SYSTEMS ANALYSIS OF CORE ARCHITECTURES REGULATING CELLULAR RESPONSES UNDER STRESS IN EUKARYOTES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Anirikh Chakrabarti

January 2012

SYSTEMS ANALYSIS OF CORE ARCHITECTURES REGULATING CELLULAR RESPONSES UNDER STRESS IN EUKARYOTES

Anirikh Chakrabarti, Ph.D.

Cornell University 2012

Eukaryotes have developed evolutionarily conserved mechanisms to respond to diverse ranges of internal and external perturbations e.g., changes in oxygen/nutrient levels, temperature oscillations, protein folding load, viral/bacterial attacks and graft implants. Alterations and malfunctions within these core regulatory architectures provide a peek into shifts towards cancerous and disease states. These core architectures include signal integrators and actuators, positive and negative feedback mechanisms which mediate the effectiveness and ultimate outcome of the response. Contemporary modeling approaches in the era of genomics revolution and high-throughput technology presents a unique systems level insight into many areas of biology, ecology, developmental biology and immunology. In my research, I have used a combination of *bottom-up* analysis (signaling networks based analysis) and a *top-down* analysis (using microarray/high throughput experimental data) to investigate stress responses in eukaryotes. Using the *bottom-up* analysis scheme, we assembled a series of molecular modules describing different aspects of the cellular response to stress. Some of these modules included the Unfolded Protein Response (UPR), Hypoxic Response (HR) and Tumor Angiogenesis, Epithelial to Mesenchymal Transition (EMT), Translation Initiation and Renal Allograft Failure. These modules were modeled using mass action kinetics with an Ordinary Differential Equation (ODE) based framework to investigate the internal regula-

tory cores. For example, in UPR we identified the differential negative feedback of activating transcription factor 4 (ATF4) as the key in the adaptation phase via regulation of binding immunoglobulin protein (BiP). Similarly in HR, we identified the role of activator protein 1 (AP1) in mediating the autocrine response via vascular endothelial growth factor (VEGF) and interleukin 8 (IL8) signaling modules. Model generation was done using UNIVERSAL, an in house software freely available at google code. We addressed issues pertaining to uncertainties within these models by developing POETs, a multi-objective optimizing algorithm which allowed us to train our models with experimental data from the literature. POETs presented us with an advantage by generating an ensemble of models consistent with experimental data. The diversity within these ensembles were used to study different operational paradigms within these modules. For example, in EMT we identified the differential modes of crosstalk between mitogen-activated protein kinase (MAPK) and SMADs in mediating the cellular transformation. These different modes of operation suggest insights into different diseases and irregularities in cellular adaptation. We subsequently analyzed these models using parameter independent structural analysis tools like extreme pathways and parameter dependent tools such as fragility and robustness to identify targets relevant to therapeutic interventions. These configurations represent experimentally testable hypothesis and potentially new strategies to manipulate the cellular responses. At an intermediate (length-scale) level, we developed multiscale modeling strategies by inductively extrapolating the consequences of cell signaling to tissue/organ function. We employed this (signaling assisted multiscale modeling (SAMM)) strategy to investigate tumor growth and angiogenesis. Using a *top down* strategy, we used microarray datasets to investigate cellular signaling, identify malfunctions and create predictive

models to infer patient outcome in case of hypoxia induced tumor growth and angiogenesis.

BIOGRAPHICAL SKETCH

Anirikh Chakrabarti was born in Kolkata, India. His interest in science was fueled and supported by his parents (Mr. Tapas Chakrabarti and Mrs. Bithika Chakrabarti) and his cousin (Gaurav Chakravorty). He went on to complete his undergraduate Bachelor of Technology (BTech) degree in Metallurgy and Material Sciences and Master of Technology (MTech) degree in Metallurgy and Material Sciences from the Indian Institute of Technology Bombay (IITB) at Mumbai in India. Here he worked on the Modeling of Cerebrospinal Fluid Dynamics for his masters thesis under the able guidance of Dr. N. B. Ballal. Upon graduating from IITB, he joined the group of Dr. Jeffrey D. Varner at Cornell University as a graduate student. Here under the able guidance of Dr. Varner, Dr. DeLisa and Dr. Shuler he worked on a diverse set of projects under a broad classification as "Systems Analysis of Core Architectures Regulating Cellular Responses Under Stress in Eukaryotes". He also had the opportunity to be a graduate student advisor and research team leader wherein he mentored 30+ undergraduate students working on multiple projects in the VarnerLab at Cornell University. He also mentored two masters students and an extremely motivated high school student. He was also a teaching assistant in several courses ranging from *Principles of Biomolecular Engineering*, *Chemical Kinetics and Transport* to *Microchemical and Microfluidic Systems*.

To my parents, family and friends ...

ACKNOWLEDGEMENTS

I would first like to thank Dr. Jeffrey D. Varner for providing me with an opportunity to work in the *VarnerLab* and for his continued support, guidance, encouragement and motivation. Dr. Varner has been a true mentor, guide and guardian to me for the last few years. His passion towards research and attention to detail will always be a source of inspiration to me. I would also like to thank my committee members Dr. Shuler and Dr. DeLisa. for their words of advice. Their stimulating questions and suggestions have greatly contributed to improving the quality of my work. I would like to thank our collaborators; Dr. Claudia Fischbach, Dr. Abraham Stroock, Dr. Scott Verbridge and Dr. David B. Leeser, for providing an opportunity to collaborate with their laboratory. Working with them has been a wonderful learning experience and has gone a long way in shaping the direction of my research.

A special thanks to the VarnerLab members, especially Dr. Ryan Tasseff, Dr. Satyaprakash Nayak and Dr. Sang Ok Song for their support, advice and their constructive suggestions and criticisms. I would also like to thank all the undergraduates who worked with me for their efforts they put in all my projects. A special thanks to Shelby Clark-Shevaliar, Steve Thompson and Carol Casler for their support and help in making my stay in Cornell possible.

I would like to extend my gratitude towards my family and friends for their support. Their love and support in all stages of my life has been the underlying backbone of my success. Lastly, I would like to thank Cornell University, Weill Cornell Medical College and the various funding agencies for the scholarships awarded to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Eukaryotes have developed evolutionarily conserved mechanisms to respond to diverse ranges of internal and external perturbations, e.g. changes in oxygen/nutrient levels, temperature oscillations, protein folding load, viral/bacterial attacks and graft implants ([237, 492, 627, 625, 626, 628, 220, 15] to name a few). Alterations and malfunctions within these core regulatory mechanisms provide a peek into shifts towards cancerous and disease states. These core architectures include signal integrators and actuators, and positive and negative feedback mechanisms which mediate the effectiveness and ultimate outcome of the response. Contemporary modeling approaches in the era of genomics revolution and high-throughput technology offer a unique systems level insight into many areas of biology, ecology, developmental biology and immunology [327]. Systemic understanding of disease initiation and progression, by incorporating conventional wet experiments and modeling to predict and optimize therapies over a wide range of length scales, could help identify strategies for personalized medicine [100]. Computational systems biology presents us with tools and strategies to do just this [258, 257, 141]. It is poised to perform large number of *in-silico* experiments, faithfully which otherwise would have been very time/money/resources consuming. Also, using tools of knowledge discovery and data-mining it can be used to guide experiments and obtain valuable insight into biological systems.

However, biological systems involve intricate symbiotic/asymbiotic relationships and crosstalk between several length scales. Biological complexity is an exponential function of number of components and interactions, thereby

escalates rapidly with each increment in level of organization [59]. This opens up several directions of research commonly aiming at probing and understanding biological complexity, ultimately influencing drug discovery. In a *bottom-up* approach, systems biology research focusses on understanding/investigating physiology and disease from a level of molecular pathways and regulatory pathways [59] (Fig. 1.1). From targeting specific areas of cellular signaling, e.g. modeling apoptosis [286, 285] to whole disease states, e.g. breast cancer [425, 272, 18]; this approach is widely being used. Hyduke *et. al.* are even working towards reconstructing genome-scale signaling networks [221]. While attractive, this area comes with several inherent challenges as well.

In a *top-down* approach, the attempt is to reverse-engineer integrative models of human physiology and disease using high-throughput data. Broadly, this task could be divided to two sub-tasks: extract relevant sub-data and infer connectivity. To address the area of inferring connectivity from such high-throughput data, several groups are working to identify strategies for the same [136, 52, 51, 96, 95, 503, 594, 528, 169]. Here size is a big issue. Researchers use clustering or correlation based techniques to reduce the complexity of these data sets by identifying clusters of components [528, 169, 303, 594, 24, 293]. However these techniques are insufficient to determine the kinetic or dynamic parameters required to model the dynamical behavior of the system [508]. Further mathematical mining tools are also being investigated to infer valuable information from these data-sets, e.g. SVD analysis [568, 252, 87, 632, 607, 530, 8, 9, 104]. Even with these constraints, this approach has been successfully used to investigate several disease states, e.g. B cell lymphoma [5], breast tumors [426, 419]. An intermediate yet extremely challenging sub-area of work is that of multiscale modeling involving extrapolation of the molecular signaling events to tissue-

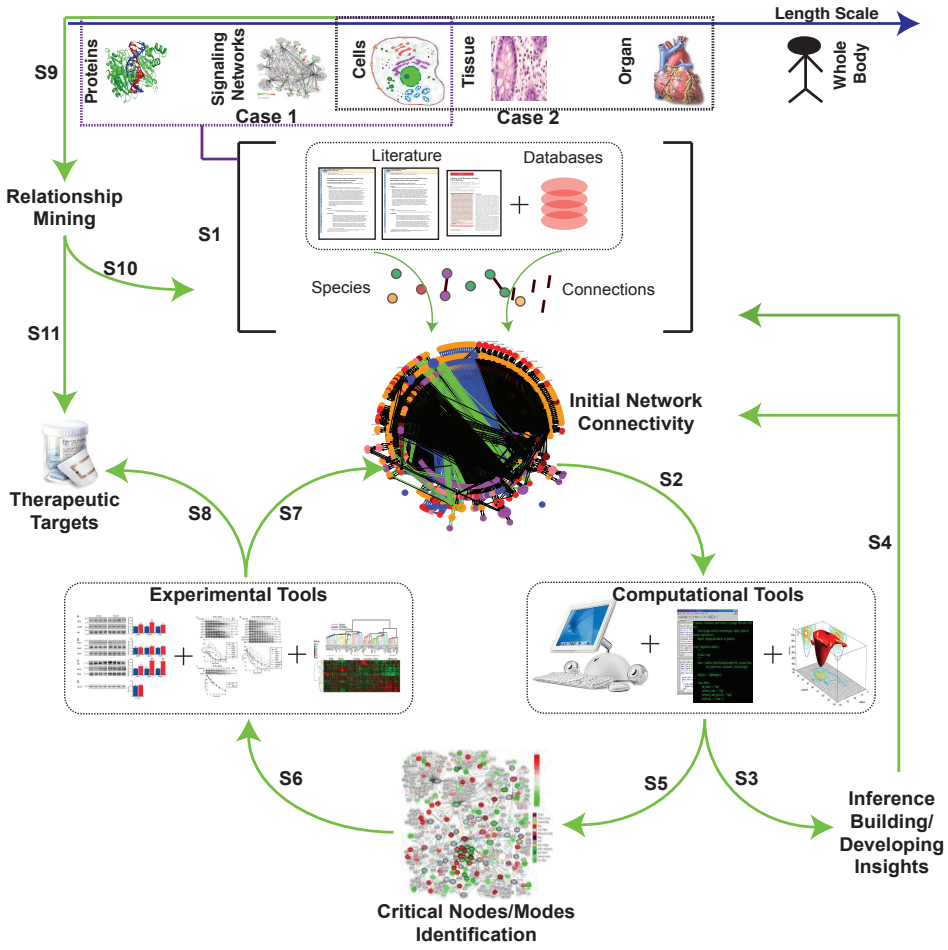


Figure 1.1: Schematic of the strategies explored in the current study: A combination of *bottom-up* analysis (signaling networks based analysis) and a *top-down* analysis (using microarray/high throughput experimental data) was used to investigate stress responses in eukaryotes. Using steps S1-S4 we can explain complex behaviors in biological systems. Using steps S1-S8, we can explore and identify fragile/robust features of complex signaling networks and identify therapeutically relevant targets. Using S9-S11, we can mine high throughput data (e.g., microarray data) for functional relationships and new insights into cellular functioning.

level/organ-level phenotypes. While cellular networks provide deductive capabilities in investigating biological mechanisms, there is a need to inductively extrapolate the consequences of cell signaling to tissue/organ function. This presents an emerging need of models taking both signaling networks and different biochemical and biophysical contexts into account [11, 310, 326, 434, 578]. Computational multiscale modeling approaches are poised to address these growing needs [578].

In my research, I have used a combination of *bottom-up* analysis (signaling networks based analysis) and a *top-down* analysis (using microarray/high throughput experimental data) to investigate stress responses in eukaryotes. This form of research is by nature interdisciplinary and involved the participation of experimentalists and fellow researchers from a wide range of academic disciplines (please see acknowledgements). Broadly, I shall classify my work into the three strategical bins as discussed below.

1.1 *Bottom up* analysis of large complex cellular programs: Modeling signaling networks.

Generally computational models used to investigate cellular signaling networks are those with biochemical/biophysical mechanistic detail [256, 600, 59, 286, 285, 425, 272, 18, 526, 320, 319, 369]. This *bottom-up* approach makes use of laws of mass action and/or Michaelis - Menten enzyme kinetics to mechanistically represent individual biochemical reactions in a signaling network. These models are definitely viable/useful in mining/elucidating relationships between the participating components [4]. The primary assumption here is that the individ-

ual biochemical reactions represented are sufficient to describe the overall signaling network dynamics. However, predictions/analysis based on this form of modeling is usually sensitive to missing/incorrect species, interactions or both [526, 320, 319, 369]. Additionally these models require a large number of parameters (rate constants, initial conditions) to be appropriately constrained. It is typically impossible to identify uniquely model parameters, even with extensive training data and perfect models [148]. These bring forward the development of parameter estimation and ensemble routes of analysis of such systems [26, 280, 54, 401, 505]. It also highlights the utility and relatively open area of parameter independent routes of development and analysis of such models [463, 186, 589, 57, 158, 406, 404, 415, 253]. Even with these limitations, the detailed yet simplified representation of the biochemical and biophysical mechanisms enables the ease of model generation using softwares [487, 44, 561, 79] and to perform predictive computational experiments, which can later be validated experimentally.

Bhalla and Iyengar pioneered the use of large scale signaling network models, integrating several signaling pathways to study neurons [39]. The authors demonstrated how crosstalk between signaling pathways could give rise to emergent behaviors such as signal integration across time scales, bistability, and feedback. This form of investigation has made immense progress over the years leading to several prominent studies [600, 59, 286, 285, 425, 272, 18, 526, 320, 319, 369]. Mechanistic models are also useful for understanding the fundamental design principles underlying biological networks. Alon and colleagues combined modeling and experimental studies to demonstrate how common network motifs found in signaling networks can give rise to a diverse spectrum of systems properties such as network robustness, signaling acceleration/deceleration, and

memory [7, 353]. Taken together, biochemically and biophysically mechanistic models of cell signaling are useful for understanding the mechanisms for cell signaling regulation and the relationships between cell signaling networks and the functions they regulate.

By enabling the identification of robust/fragile parts of signaling networks, computational modeling has led to new ways of rationally designing multi-target or multi-component therapies. With this idea in mind, we assembled a series of molecular modules describing different aspects of the cellular response to stress. Some of these modules included the Unfolded Protein Response (UPR), Hypoxic Response (HR) and Tumor Angiogenesis, Epithelial to Mesenchymal Transition (EMT), Insulin mediated Translation Initiation and Renal Allograft Failure. These modules were modeled using mass action kinetics with an Ordinary Differential Equation (ODE) based framework to investigate the internal regulatory cores. For example, in UPR we identified the differential negative feedback of ATF4 as the key in the adaptation phase via regulation of BiP (details in Chapter 3). Similarly in HR, we identified the role of AP1 in mediating the autocrine response via VEGF and IL8 signaling modules (details in Chapter 6). Model generation was done using UNIVERSAL, an in house software freely available at google code. We addressed issues pertaining to uncertainties within these models by developing POETs, a multi-objective optimizing algorithm which allowed us to train our models with experimental data from the literature [504]. POETs presented us with an advantage by generating an ensemble of models consistent with experimental data. The diversity within these ensembles were used to study different operational paradigms within these modules. For example, in EMT we identified the differential modes of crosstalk between MAPK and SMADs in mediating the cellular transformation

(details in Chapter 5). These different modes of operation suggest insights into different diseases and irregularities in cellular adaptation. We subsequently analyzed these models using parameter independent structural analysis tools like extreme pathways and parameter dependent tools like fragility and robustness to identify targets relevant to therapeutic interventions. These configurations represent experimentally testable hypothesis and potentially new strategies to manipulate the cellular responses.

1.2 Modeling integration of intracellular signaling to tissue/organ level behavior: Multiscale modeling.

At the molecular level, signaling networks implicated in several disease states exhibit considerable complexity in size and connectivity. The intrinsic complexity of these networks complicates the interpretation of experimental findings. Added to this is the dynamical interactions between the cells and their surrounding microenvironmental changes. Extrapolating the molecular signaling events to tissue-level/organ-level phenotypes adds another layer of complexity across spatio-temporal and functional scales. While cellular networks provide deductive capabilities in investigating biological mechanisms, there is a need to inductively extrapolate the consequences of cell signaling to tissue/organ function. This presents an emerging opportunity and the need for models taking both signaling networks and different biochemical and biophysical contexts into account [11, 310, 326, 434, 578]. Computational multiscale modeling approaches are poised to address these growing needs [578]. Multiscale models have been used extensively in several disease states e.g. in cardiac physiol-

ogy [600], epidermal wound healing [514], renal physiology [112] to identify biological mechanisms, infer multiscale consequences to cell signaling activities. Multiscale modeling as such is modular in nature and involves integrating common linkages (species/components) across individual computational models of different length scales. For example, Cortassa *et. al.* developed an integrated model of the guinea pig cardiomyocyte, linking cell electrophysiology, force generation, and mitochondrial energy generation to investigate phenomena such as oxidative-stress induced action potential shortening [84].

Disease states like cancer are heterogeneous cellular entities involving dynamical interactions between genetically altered cells and their surrounding microenvironmental changes. These forms of disease states present an emerging importance and need of models taking both signaling networks and different biochemical and biophysical contexts into account. For example, factors like vascular endothelial growth factor (VEGF) and/or interleukin-8 (IL8) have been implicated in multiple studies in the regulation of some tumor growth as well as in the activation and chemotactic guidance of capillary sprouting and growth. Detailed computational models of such complex phenomena using intricate cellular signaling mechanisms, nutrient/growth factor transport, movement/interactions between normal and tumor cells is critical in identifying the key regulators of such a system. We have created a signaling assisted multiscale modeling approach (SAMM) to study systems across multiple length scales. Key highlights of SAMM include use of detailed cellular models to regulate cellular decisions, use of an ensemble of models to incorporate differences between cells, coupling intracellular markers and probabilistic rules to monitor cellular fate and cellular motility and a hybrid approach to improve speed of solving of external and internal mass-balances. SAMM allows us the flexibility to in-

clude multiple cell types, freedom of seeding cells with specific internal cellular machinery within complex geometries, regulate multiple cellular decisions using internal and secreted cellular markers, ease of computation and improved solving speed.

We demonstrated using SAMM the ability to capture and identify spatio-temporal effects upon growth, ability to model and monitor effects of internal variations in the cells e.g., species knockouts (KO's) and ultimately ability to monitor effects of coupling probabilistic rules to internal marker levels and its effects on growth morphology and growth dynamics. For this study we used a tumor/endothelial cell model which has been used earlier to study eukaryotic hypoxic response [70]. Both VEGF and IL8 have been shown to regulate, in an autocrine signaling based mechanism, the ability of the tumor to grow. While this was used to investigate tumor growth, a similar set-up with different cell-types could be used to study other phenomenon e.g. angiogenesis, transplant failure. There are however several open areas of development in this form of modeling, e.g. using further detailed/comprehensive cellular models, efficient ways of solving extracellular/intracellular mass balances, incorporating stress and extracellular matrix properties, improved models of cellular fate/decisions based on combination of internal and external species/conditions, cell size variations, flow and other considerations.

1.3 *Top down* analysis: High-throughput data coupled with physiochemical model based assessment of biomarker importance.

Recent advancements in high-throughput methods for characterizing genomic, transcriptomic, proteomic, and metabolomic states allow one to view the global consequences of molecular perturbations rather than just the *usual suspects* [508]. However, this wealth of *high-throughput* data creates new challenges in data interpretation, because most of the measurements lack a biological context for interpreting the biological relevance to the experimental perturbation. These measurements give a snapshot of activity at various levels at the time of isolation [508, 606, 166]. While this has successfully provided a detailed view in several disease states, e.g. B cell lymphoma [5], breast tumors [426, 419]; it lacks providing us with the understanding of relationships within gene networks [606, 166].

Interactions between gene products are extremely complex and are not the only determinant of cellular behavior [508]. To address the area of inferring connectivity from such high-throughput data, several groups are working to set up tools [136, 52, 51, 96, 95, 503, 594, 528, 169]. Researchers use clustering or correlation based techniques to infer information about the topology of the connections extensively [528, 169, 303, 594, 24, 293]. These techniques help reduce the complexity of these data sets by identifying clusters of signaling species that may either be co-regulated or that can similarly regulate other species in a signaling network. However these techniques are insufficient to determine the kinetic or dynamic parameters required to model the dynamical behavior of the system [508]. Another strategy in extracting potential gene relationships from

such high-throughput data is by using singular value decomposition (SVD) [568, 252, 87, 632, 607, 530, 8, 9, 104]. These approaches draw on the use of computational tools to identify features in the data that may globally represent the entire data set (e.g., principal components). These approaches inherently make few/no assumptions about the data and have the capability to provide unbiased identification of unexpected correlations. These correlations might not explain the causality between correlated components, however they can be useful for identifying nonintuitive patterns in the data and help guide future experiments.

In the current study, using a *top down* strategy, we used these extensive datasets to investigate cellular signaling, identify malfunctions, create predictive models to infer patient outcome. Specifically we employed this strategy in case of hypoxia induced tumor growth and angiogenesis. We created an in house platform to analyze microarray data and overlay the quantitative information on pre-curated pathways to identify malfunctions within the cells. We coupled this to mathematical analysis techniques such as singular value decomposition (SVD) to identify key mediators of the response. Using these strategies, we identified differential regulators of the cellular behavior in 2D versus 3D experimental protocols. We also identified differential behaviors and involvement of key signaling pathways in mediating the hypoxic response. While this form of analysis provided key insights in the experimental directions, which are being pursued by our collaborators, we specifically used this insight to create proof-of-concept physiochemical models to further the process of modeling these disease states. This strategy has immense scope of development, from identifying key regulators to forming proof-of-concept physiochemical models.

CHAPTER 2

A REVIEW OF THE MAMMALIAN UNFOLDED PROTEIN RESPONSE

Authors - Chakrabarti A, Chen AW, Varner JD.

Published in - Biotechnol Bioeng. 2011 Dec; 108(12):2777-93.

2.1 Abstract

Proteins requiring post-translational modifications such as N-linked glycosylation are processed in the endoplasmic reticulum (ER). A diverse array of cellular stresses can lead to dysfunction of the ER and ultimately to an imbalance between protein-folding capacity and protein-folding load. Cells monitor protein folding by an inbuilt quality control system involving both the ER and the Golgi apparatus. Unfolded or misfolded proteins are tagged for degradation via ER associated degradation (ERAD) or sent back through the folding cycle. Continued accumulation of incorrectly folded proteins can also trigger the Unfolded Protein Response (UPR). In mammalian cells, UPR is a complex signaling program mediated by three ER transmembrane receptors: activating transcription factor 6 (ATF6), inositol requiring kinase 1 (IRE1) and double-stranded RNA-activated protein kinase (PKR)-like endoplasmic reticulum kinase (PERK). UPR performs three functions, adaptation, alarm and apoptosis. During adaptation, the UPR tries to reestablish folding homeostasis by inducing the expression of chaperones that enhance protein folding. Simultaneously, global translation is attenuated to reduce the ER folding load while the degradation rate of unfolded proteins is increased. If these steps fail, the UPR induces a cellular alarm and mitochondrial mediated apoptosis program. UPR malfunctions have been associ-

ated with a wide range of disease states including tumor progression, diabetes, as well as immune and inflammatory disorders. This review describes recent advances in understanding the molecular structure of UPR in mammalian cells, its functional role in cellular stress, and its pathophysiology.

2.2 Introduction

Protein folding is strategically important to cellular function. Secreted, membrane-bound and organelle-targeted proteins are typically processed and folded in the endoplasmic reticulum (ER) in eukaryotes [363, 456, 245]. Intracellular perturbations caused by a variety of stressors disturb the specialized environment of the ER leading to the accumulation of unfolded proteins [117, 134]. Glucose deprivation, aberrant calcium regulation, viral infection and hypoxia can all alter protein folding and induce ER stress [456, 245]. Physiological processes such as aging can also influence protein folding [363]. Normally, cells ensure that proteins are correctly folded using a combination of molecular chaperones, foldases and lectins [363]. However, when proper folding can not be restored, incorrectly folded proteins are targeted to ER Associated Degradation (ERAD) pathways for processing [245]. If unfolded or misfolded proteins continue to accumulate, eukaryotes induce the unfolded protein response (UPR).

In mammalian cells, UPR is a complex signaling program mediated by three ER transmembrane receptors: activating transcription factor 6 (ATF6), inositol requiring kinase 1 (IRE1) and double-stranded RNA-activated protein kinase (PKR)-like endoplasmic reticulum kinase (PERK). UPR performs three functions, adaptation, alarm and apoptosis. During adaptation, the UPR tries to

reestablish folding homeostasis by inducing the expression of chaperones that enhance protein folding. Simultaneously, translation is globally attenuated to reduce the ER folding load while the degradation of unfolded proteins is increased. If these steps fail, the UPR induces a cellular alarm and apoptosis program. The alarm phase involves several signal transduction events, ultimately leading to the removal of the translational block and the down-regulation of the expression and activity of pro-survival factors such as the B-cell lymphoma 2 (Bcl2) protein. After the alarm phase, cells can undergo apoptosis, although ER stress can also initiate autophagy [386, 609, 36, 240, 218, 270, 146]. Thus, ER folding homeostasis strongly influences physiology [134]. Aberrant protein folding and UPR have been implicated in a number of pathologies. For example, the onset of diabetes [478] as well as myocardial ischaemia, cardiac hypertrophy, atherosclerosis and heart failure [164] have all been linked with aberrant folding or UPR signaling. This review describes the molecular basis of UPR in mammalian cells, and its functional role in cellular stress induced pathophysiology. First, we explore the regulation and activity of the IRE1, ATF6 and PERK stress receptors. Next, we review the role of aberrant UPR and ER-stress in a variety of diseases, including cancer and neurodegenerative disorders. Lastly, we present recent strategies to manipulate UPR which have been shown to enhance protein titers in mammalian hosts.

2.2.1 The folding cycle, quality control and ER associated degradation (ERAD).

Newly synthesized polypeptide chains enter the ER through a peptide translocon in the ER membrane composed of four proteins, Sec61α,β,γ and TRAM [340]. Upon entering the ER, these nascent chains begin to fold, often as they are being co-translationally modified [124]. In the cytosol, protein folding is largely driven by the collapse of hydrophobic side chains which ultimately form the core of the folded protein. Another key factor, the burial of electrostatic interactions, e.g., salt brides or hydrogen bonds in the hydrophobic core, constrains possible folding choices in a very complex free energy landscape. The forces that govern cytosolic folding are also active in the ER. However, protein folding in the ER is more complex because of post-translational modifications such as disulfide bond formation or N-linked glycosylation. Interestingly, the folding landscape in the ER can sometimes be traversed quickly; small proteins like the Semliki Forest virus capsid protein can fold in approximately 50ms [470]. However, for more complex proteins, e.g., the human coagulation factors V and VIII, folding can take several minutes to hours to complete [424].

The folding quality of proteins in the ER is maintained by an in-built quality control (QC) system (Fig. 2.1) which ensures proteins are in their native folded state before exiting the ER [114, 117]. A protein is correctly folded if it has attained its native conformation after required co- or post-translational modifications. On the other hand, exposed hydrophobic regions, unpaired cysteine residues, or aggregation are all markers of an unfolded or misfolded conformation [114]. The best characterized QC system in the ER is the so-called *glycan-code* [195]. Most polypeptides entering the ER are modified by adding

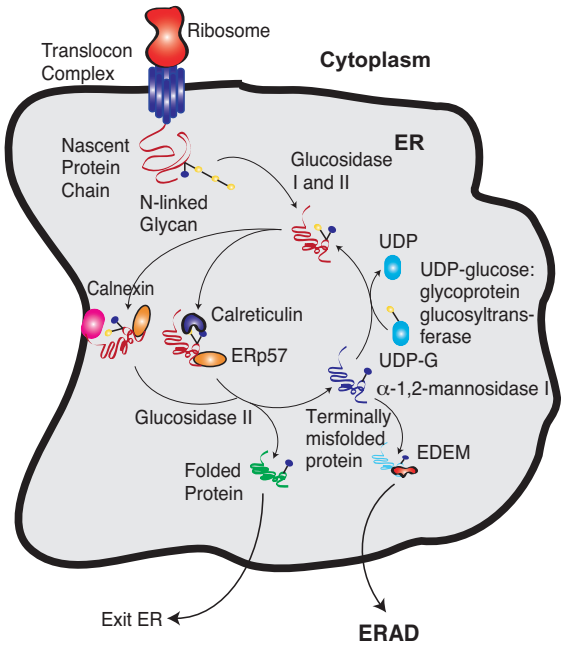


Figure 2.1: The calnexin/calreticulin protein folding cycle. Yellow circles denote glucose groups and while blue circles denote mannose groups. After entering the ER lumen, glucosidase I and II remove two glucose groups. The monoglucosylated glycoprotein then interacts with calnexin/calreticulin. These chaperones interact with the thiol-disulphide oxidoreductase ERp57. Cleavage of the last glucose residue by glucosidase II leads to the release of the chaperones. At this time, the protein could have either folded and left the ER or it could have attained an incorrect state. The incorrectly folded proteins are then the substrates of UDP glucose:glycoprotein glucosyltransferase, which puts a glucose residue back to the incorrectly folded protein. This enables the protein to spend some more time in folding in the ER. If the protein fails to fold in a repeated number of cycles, the mannose residue is removed by α-1,2-mannosidase I. This enables the protein to be recognized by ER-degradation-enhancing 1,2-mannosidase-like protein (EDEM). This targets the unfolded proteins for ER-associated degradation (ERAD).

preassembled oligosaccharides to asparagine side chains appearing in ASN-X-SER/THR motifs [199]. Once attached, these oligosaccharide groups can be sequentially modified by glucosidases I and II to form mono-glucosylated intermediates that are recognized by the ER lectins calnexin and calreticulin [180] and the associated oxidoreductase ERp57 [393, 392, 579, 313, 192]. Calnexin is a type I membrane protein with a β-sandwich carbohydrate binding domain and a hairpin, called the P-domain, extending away from the carbohydrate binding domain [480]. Calrecticulin is the soluble paralog of calnexin with only minor structural differences in the P-domain [480, 116, 115]. Calnexin and calrecticulin binding increases the efficiency of glycoprotein folding by protecting against aggregation [193] and ensures that misfolded proteins are retained in the ER [436]. These ER glycoprotein chaperones also promote disulfide bond formation through their interaction with ERp57; ERp57 binds the P-domain of both calnexin and calrecticulin thereby promoting disulfide bond formation at specific glycoprotein locations [393, 392, 142]. Glycoproteins are released from calnexin and calrecticulin by further glucose cleavage by glucosidases II. Once released, these proteins can fold into their native conformation, they can be re-glucosylated and re-processed by the calnexin/calreticulin cycle, or they can be targeted for ER-associated degradation (ERAD) (recently reviewed in Maattanen *et. al.* [325]). Interestingly, while calnexin and calrecticulin ultimately promote folding, they do not recognize misfolded proteins. Misfolded proteins can be targeted (or re-targeted) to the calnexin/calrecticulin cycle by glucosylation with UDP-glucose; the C-terminal domain of glycoprotein glucosyltransferase (GT) glucosylates near-native conformations by using its N-terminal protein sensor to locate exposed hydrophobic residues [507, 545, 63, 88]. In this way, GT acts as an adapter allowing attachment of calnexin/calreticulin and re-

processing in the calnexin/calreticulin cycle [411]. Repeated glycosylation and deglycosylation cycles ensure misfolded glycoproteins spend sufficient time in the ER to correctly fold. Exit from the calnexin/calreticulin cycle can also lead to ERAD. In this case, terminal mannose residues are removed from the attached oligosaccharides by ER α-mannosidase I, which leads to interaction with membrane localized ER degradation enhancing α-mannosidase like proteins (EDEM and EDEM2) [217, 391] and subsequent retro-translocation to the cytosol. The retro-translocon responsible for transporting unfolded or misfolded proteins out of the ER is uncertain [546]; several ER membrane proteins have been proposed including Sec61 components, Derlin family members and E3-ubiquitin ligases (reviewed in [194]). Mannose removal decreases the likelihood that a unfolded protein will be processed in the calnexin/calreticulin cycle [114], thus, increasing the probability of terminal mannose cleavage and retro-translocation. Once in the cytosol, these unfolded or misfolded proteins are degraded by the ubiquitin proteasome system [205].

Hydrophobic unfolded or misfolded proteins are recognized in the ER by molecular chaperones which bind these proteins and increase the probability of correct folding [138, 198, 200]. Similar to calnexin and calrecticulin, hydrophobic chaperone binding increases the residence time of unfolded proteins in the ER lumen, giving these proteins a chance to fold (Fig. 2.2A). For example, the HSP70 family of chaperones recognize, in an ATP-dependent manner, exposed hydrophobic patches on a broad spectrum of unfolded or misfolded proteins [245]. Repeated binding and release of HSP70 chaperones ensures that incorrectly folded proteins do not exit the ER [245]. One critical member of the HSP70 family is BiP or GRP78. BiP consists of an N-terminal ATPase domain and a C-terminal peptide binding domain [157]. When bound to ATP, BiP binds

unfolded hydrophobic stretches with low-affinity. However, unfolded or misfolded protein binding stimulates the N-terminal ATPase activity of BiP resulting in an ADP-bound form with much higher affinity for hydrophobic motifs [157]. Interestingly, the affinity of BiP (and other HSP70 family members) for ADP is approximately six-fold greater than for ATP. Thus, nucleotide exchange factors (NEFs) such as BiP associated protein (BAP), Sil1/S1s1p or GrpE-like proteins are required to catalyze the ADP/ATP exchange needed for the dissociation of BiP from unfolded proteins [547]. In addition to its role as a folding chaperone, BiP also functions as an ER stress regulator by buffering Ca^{2+} levels [447]. BiP interacts with ER-localized caspase-7 [447] and prevents the activation of pro-apoptotic Bcl2 family members such as Bax [437, 144]. Beyond these activities, BiP also regulates the activation of the three transmembrane ER stress transducers: PERK, ATF6 and IRE1. Normally, BiP is bound to these ER receptors, blocking their activation. However, in the presence of exposed hydrophobic residues BiP disassociates, allowing PERK, ATF6 and IRE1 activation (Fig. 2.2B). Overexpression of BiP leads to reduced activation of IRE1 and PERK [38, 262]. The PERK and ATF6 branches are thought to be activated before IRE1 [515]; this ordering is consistent with the signals that each branch transduces. The PERK and ATF6 pathways largely promote ER adaptation to misfolding, while IRE1 has a dual role, transmitting both survival and pro-apoptotic signals.

2.2.2 Double-stranded RNA-activated protein kinase (PKR)-like endoplasmic reticulum kinase (PERK) pathway:

The PERK branch of UPR transduces both pro-survival as well as pro-apoptotic signals following the accumulation of unfolded or misfolded protein in the ER. However, its main function is to modulate translation. PERK is a type I transmembrane protein, composed of a ER luminal stress sensor and a cytosolic protein kinase domain. Dissociation of BiP from the N-terminus of PERK initiates dimerization and autophosphorylation of the kinase domain at T981 [246]. The C-terminal kinase domain shares homology with the eukaryotic translation initiation factor 2α (eIF2α) and the kinases, protein kinase R (PKR), heme-regulated inhibitor (HRI) kinase, and the gene control non-derepressible-2 (GCN2) kinase [184, 494, 493]. The eIF2α protein, which is composed of three subunits, is critical to translation initiation in eukaryotes, including GTP-dependent start-site recognition [351]. Phosphorylation of the α subunit of eIF2α blocks the exchange of GDP bound to eIF2α; hence, eIF2α remains bound to partner initiation factors (eIF2B) and translation initiation is blocked [156]. Activated PERK can phosphorylate eIF2α at S51 [184, 444], which leads to three downstream effects. First, phosphorylated eIF2α globally attenuates translation initiation. Decreased translation reduces the influx of protein into the ER, hence diminishing the folding load. Translation attenuation is followed by increased clearance of the accumulated proteins from the ER by ERAD and expression of pro-survival genes. For example, PERK activation induces of the expression of cellular inhibitor of apoptosis (cIAP) [179]. Interestingly, decreased protein translation is not universal; genes with internal ribosome entry site (IRES) sequences in the 5′ untranslated regions bypass the eIF2α translational block [481]. One of the most

well-studied of these, *ATF4*, encodes a cAMP response element-binding transcription factor (C/EBP) [318] ATF4 that drives the expression of pro-survival functions such as amino acid transport and synthesis, redox reactions and protein secretion [185]. Taken together, these effects seem to be largely pro-survival. However, ATF4 can also induce the expression of pro-apoptotic factors. For example, ATF4 induces the expression of the transcription factor C/EBP homologous protein (CHOP), which is associated with apoptotic cell-death. CHOP (also known as GADD153) is 29 kDa protein composed of an N-terminal transcriptional activation domain and a C-terminal basic-leucine zipper (bZIP) domain that is normally present at low levels in mammalian cells [455]. The transcriptional activator domain is positively regulated by phosphorylation at S78 and S81 by p38 MAPK family members [574, 344] while the bZIP domain plays a key role in the homodimerization of the protein [341, 344]. CHOP activity promotes apoptosis primarily by repression of Bcl2 expression and the sensitization of cells to ER-stress inducing agents [167, 345]. For example, Matsumoto *et al.* showed that ectopic expression of CHOP in M1 myeloblastic leukemia cells reduced Bcl2 protein concentrations, while Bax levels remained constant [341]. They further established a link between CHOP expression and apoptosis in these cells. However, while CHOP expression is negatively correlated with Bcl2 levels, there is no CHOP binding site in the *bcl2* promoter [345]. McCullough *et al.* have suggested that the bZIP domain of CHOP could act with other bZIP transcription factors to regulate *bcl2* expression [345]. Thus, it's likely that the connection between CHOP expression and apoptosis is more complex than simple down-regulation of Bcl2 expression.

Given its central role in translation attenuation, cells have evolved multiple axes to regulate PERK activity. First, the cytosolic kinase domain of PERK can be

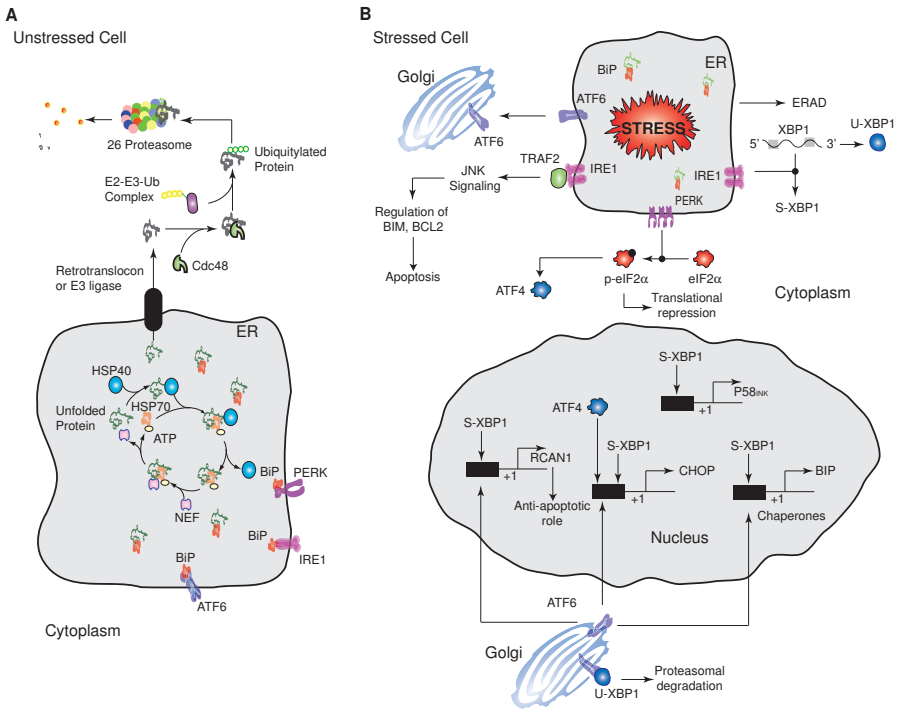


Figure 2.2: In response to the folding imbalance, cells initiate the cytoprotective unfolded protein response (UPR). **A:** The problem of unfolded or misfolded proteins in the ER is addressed by increasing the folding capacity through the up-regulation of the expression of chaperone proteins, attenuating translation by regulating eIF2α, and promoting the degradation of misfolded proteins through ER-associated degradation (ERAD). If UPR is unable to restore the folding balance, ER stress will eventually lead to apoptotic cell-death. **B:** The three signal transduction pathways mediating the unfolded protein response in higher eukaryotes. First, the PRKR-like ER kinase (PERK) pathway is initiated after BiP dissociation from PERK. While PERK transduces both pro- and anti-apoptotic signals, its main function is translation attenuation through the phosphorylation of eIF2α. Next, the activating transcription factor 6 (ATF6) pathway is activated following BiP dissociation. ATF6 induces the expression of chaperones e.g., BiP as well as apoptosis effectors such as CHOP. Lastly, the inositol-requiring kinase 1 (IRE1) pathway is activated following BiP dissociation from IRE1. Activated IRE1 has both an endoribonuclease and a serine-threonine kinase activity that drive can pro-apoptotic signals.

inhibited by the action of the DNAJ family member P58^{IPK}. P58^{IPK} was initially discovered as an inhibitor of the eIF2α protein kinase PKR [292]. P58^{IPK}, whose expression is induced following ATF6 activation, binds to the cytosolic kinase domain of PERK, inhibiting its activity [599, 558]. Inhibition of PERK kinase activity relieves eIF2α phosphorylation, thereby removing the translational block. Interestingly, P58^{IPK} expression occurs several hours after PERK activation and eIF2α phosphorylation. Thus, P58^{IPK} induction may mark the end of UPR adaptation, and the beginning of the alarm/apoptosis phase of the response [515]. In addition to its direct interaction with PERK, P58^{IPK} is also involved in co-translational protein degradation [398]. In this role, P58^{IPK} is thought to recruit Hsp70 to the cytosolic opening of the ER translocon in an effort to extract stalled nascent proteins. Interestingly, P58^{IPK} has also been identified as an ER-luminal co-chaperone acting in conjunction with BiP, although BiP-independent association between P58^{IPK} and a mutant vesicular stomatitis virus envelope glycoprotein (VSV-G^{ts045}) was also observed [422]. The activity of P58^{IPK} appears to be regulated; recently, Ni *et al.* reported a novel UPR-inducible cytosolic BiP isoform (GRP78va), generated by alternative splicing and IRES mediated translation, that antagonizes cytosolic P58^{IPK} in several human and mouse cell-lines [375]. Second, PERK induces a negative feedback loop, through its downstream effector CHOP, involving the direct de-phosphorylation of eIF2α. CHOP induces the expression of GADD34 which, in conjunction with protein phosphatase 1 (PP1), assembles into a phosphatase which dephosphorylates the S51 residue of eIF2α [384]. GADD34 is a member of the GADD family of genes which are induced by DNA damage and a variety of other cellular stresses [621]. The GADD34 binding partner in this complex appears to be responsible for PP1α recognition and targeting of the phosphatase complex to the ER. As-

sociation between GADD34 and PP1 is encoded by a C-terminal canonical PP1 binding motif, KVRF, while approximately 180 residues, near the N-terminus of GADD34, appear to be responsible for ER localization [56].

2.2.3 Activating transcription factor 6 (ATF6) pathway:

ATF6 activation involves a complex series of translocation and irreversible proteolytic processing steps, ultimately leading to the up-regulation of a pro-survival transcriptional program, in the presence of unfolded or misfolded proteins. First, ATF6 is translocated from the ER to the golgi apparatus following BiP dissociation. Once resident in the golgi, ATF6 is proteolytically processed, after which it drives its transcriptional program. ATF6 is a 90 kDa ER transmembrane protein with two homologs: ATF6α [191, 178] and ATF6β [354, 249, 190]. The ATF6α homolog is thought to be primarily responsible for transcriptional regulation of pro-survival genes following ER stress, however, ATF6β may also play a role [190, 598, 537]. Similar to IRE1 and PERK, ER stress leads to the dissociation of BIP from the N-terminus of ATF6, followed by translocation and activation. However, the mechanism of BiP interaction with ATF6 and the factors controlling ATF6 translocation to the golgi remain uncertain. N-terminal golgi localization sequences (GLS1 and GLS2) seem to be involved with BiP regulation of ATF6. BiP binding to the N-terminal GLS1 promotes the retention of ATF6 in the ER [489]. On the other hand, the GLS2 domain was required to target ATF6 to the golgi body following BiP dissociation from GLS1 [489]. The lectin CRT might also have a role in keeping the ATF6 in the ER [215]. Unlike the previous two kinase pathways, ATF6 activation does not involve phosphorylation of a C-terminal kinase domain. Rather, after translocated to

the golgi, ATF6 undergoes regulated intramembrane proteolysis (RIP); the luminal domain is first cleaved by serine protease site-1 protease (S1P) followed by metalloprotease site-2 protease (S2P) cleavage [191, 605, 75, 490]. Cleavage at the juxtamembrane site allows the 50 kDa transcriptional domain of ATF6 to be translocated to the nucleus where it regulates the expression of genes with ATF/cAMP response elements (CREs) [575] and ER stress response elements (ERSE) in their promoters [610, 263]. Cleaved ATF6 induces a gene expression program, in conjunction with other bZIP transcription factors and required co-regulators, such as nuclear factor Y (NF-Y) [263, 612], that increases chaperone activity as well as the degradation of unfolded proteins [598, 590]. For example, ATF6 upregulates BiP, protein disulfide isomerase (PDI) and ER degradation-enhancing alpha-mannosidase-like protein 1 (EDEM1) expression. Additionally, ATF6 induces the expression of the X box-binding protein 1 (XBP1) which, after processing by activated IRE1α, induces the expression of chaperones as well as control elements such as P58^{INK} [611]. The ATF6-induced gene expression program is also cytoprotective. For example, ATF6 induces regulator of calcineurin 1 (RCAN1) expression [32]. RCAN1 sequesters calcineurin [32], a calcium activated protein-phosphotase B, that dephosphorylates Bcl2-antagonist of cell death (BAD) at S75 or S99 [571]. This leads to sequestering of Bcl2 by Bad, which inhibits its downstream anti-apoptotic activity [571]. Recently, a number of ATF6 homologs have been identified, e.g., OASIS, CREBH, LUMAN/CREB3, CREB4 and BBF2H7 that are processed in a similar way as ATF6, yet their function remains unknown [457]. Thus, ER-stress induced ATF6 signaling may be responsible for additional undiscovered functionality.

Currently, little is known about deactivation of ATF6. Recently, XBP1u, the unspliced form of XBP1, has been implicated as a negative regulator for ATF6

[615]. Following, the induction of ER stress, two versions of XBP1 exist: XBP1u and sXBP1 [615]. In the recovery phase following ER stress, high levels of XBP1u may play a dual role. First, XBP1u binds sXBP1, promoting complex degradation [614, 541]. Second, XBP1u can bind ATF6α rendering it more prone to proteasomal degradation [615]. Taken together, these two steps may slow the transcription of ER chaperones and ERAD components during the recovery phase following ER stress.

2.2.4 Inositol-requiring kinase 1 (IRE1) pathway:

IRE1 is the most evolutionarily conserved branch of UPR and its interactome has recently been reviewed by Hetz and Glimcher [208]. IRE1 initiates a program with both pro-survival and pro-apoptotic components in the presence of misfolded or unfolded proteins. IRE1 is a 100 kDa type I ER transmembrane protein with both an endoribonuclease and a serine-threonine kinase domain [245]. IRE1 has two homologs, IRE1α and IRE1β; IRE1α is expressed in a variety of tissues [540] while IRE1β is found only in the intestinal epithelia [540, 573]. The N-terminus of IRE1, located in the ER lumen, senses unfolded or misfolded proteins through its interaction with BiP [85, 486, 499]. There has been some controversy surrounding the mechanism of unfolding-induced BiP dissociation and the role of possible direct interaction of unfolded proteins with IRE1. Normally BiP is bound to the N-terminus of IRE1 [38, 389, 309]. However, in the presence of unfolding cues BiP dissociates and is sequestered by the unfolded or misfolded proteins [255]. IRE1 may also sense unfolded cues in a BiP independent manner through a N-terminal peptide binding domain [309, 308]. Interaction of this domain with unfolded or misfolded proteins has been suggested as a pre-

cursor to BiP dissociation. Current studies in yeast have suggested that BiP first dissociates, allowing IRE1 dimerization. This step is then followed by direct sensing of unfolded motifs which work to orient IRE1 into an active signaling cluster [254]. In either case, BiP dissociation allows IRE1 activation. IRE1 is activated by homooligomerization followed by autophosphorylation of the C-terminal kinase domain at S724 [486, 584, 583, 403]. IRE1 activation enables both its kinase and endoribonuclease activities to transduce signals simultaneously through two distinct signaling axes. The endoribonuclease activity cleaves a 26-nucleotide intron from the XBP1-mRNA [491, 611, 291] which generates a 41 kDa frameshift variant (sXBP1) that acts as a potent transcription factor. sXBP1 homodimers, along with co-regulators such as nuclear factor Y (NF-Y), regulate the expression of a variety of ER chaperones and protein degradation related genes [330, 439]. sXBP1 also upregulates the expression of $P58^{IPK}$; as discussed previously, $P58^{IPK}$ is a member of the DNAJ protien family that negatively regulates PERK activity, forming one of the many modes of crosstalk between the UPR branches [599].

The cytosolic domains of activated IRE1α transduce late-phase UPR signals. Cytosolic IRE1α dimers interact with adaptors such as tumor necrosis factor receptor-associated factor 2 (TRAF2) to drive signal-regulating kinase (ASK1) activation and then subsequently cJUN NH_2-terminal kinase (JNK) and p38MAPK activation [554]. IRE1α also modulates the activation of other kinases such as extracellular signal-regulated kinases (ERKs) as well as nuclear factor κB (NF-κB) pathways [374, 219]. However, the role of these additional effectors in UPR is not well understood. ASK1 activity is regulated by phosphorylation/dephosphorylation at several sites as well as by physical interaction with other proteins. The current model for ASK1 activation, at least for TNF-mediated

activation, involves release of ASK1 from inhibitory proteins such as 14-3-3 [377], TRAF-dependent homodimerization [168] and autophosphorylation at T845 [542]. Phosphorylation of ASK1 at S83 by Akt/PKB (protein kinase B) [250] and dephosphorylation at S845 by protein phosphatase 5 (PP5) [359] both decrease ASK1 activity. More recently, the kinase PIM1 has also been shown to phosphorylate ASK1 at S83 [172]. Interaction with CDC25A and 14-3-3 proteins have also been shown to decrease ASK1 activity [637, 630]. ASK1 phosphorylates and activates two downstream kinases, MMK4 and MMK3 which in turn activate JNK and p38 MAP kinase, respectively. JNK is activated by dual phosphorylation at T183 and Y185 by MMK4 [102]. Once activated, JNK performs a number of functions including activation of the pro-apoptotic Bim protein [294, 430] and the inhibition of Bcl2 [596]. Activated JNK activates the proapoptotic Bcl-2 family member Bim by phosphorylation at S65 [294, 430]. Bim is normally sequestered by motor complexes interacting with the cytoskeleton [431]. Following JNK-mediated phosphorylation, Bim translocates to the mitochondrial outer membrane, where it promotes cytochrome c release and caspase activation [431, 73]. Interestingly, a positive feedback loop exists between Bim and caspase 3 activation; phosphorylated Bim is a caspase 3 target, that once cleaved, is a more potent inducer of cytochrome c release [73, 82]. Recent ER stress studies in MCF-7 breast carcinoma-derived cells using thapsigargin (Tg) suggested that Bim expression was regulated by CHOP; following Tg treatment, a two-fold increase in Bim mRNA and a five-fold increase in Bim proteins were observed [432]. JNK activation also regulates the activity of anti-apoptotic protein Bcl2 [596, 582]. Active JNK1 inhibits Bcl2 via phosphorylation at sites T69, S70 and S87 [582]. In contrast, other stress induced proteins like p38 family members phosphorylate Bcl2 at S87 and T56 only [94]. Ultimately, inhibi-

tion of Bcl2 and the activation of Bim leads to BAX/BAK dependent apoptosis. Thus, signals initiated from the cytosolic kinase domain of IRE1α are largely pro-apoptotic.

IRE1α activity is regulated by several proteins, including tyrosine phosphatase 1B (PTP-1B), ASK1-interactive protein 1 (AIP1) and members of the Bcl2 protein family. PTP-1B has been implicated in a number of IRE1α signaling events. The absence of PTP-1B reduced IRE1α dependent JNK activation, XBP1 splicing and EDEM transcription in immortalized and primary mouse embryonic fibroblasts [171]. However, no physical interaction between IRE1α and PTP-1B was established. On the other hand, AIP1 physically interacts with both TRAF2 and IRE1α, suggesting a model in which AIP1 facilitates IRE1α dimerization and activation [321]. The C-terminal period-like domain (PER) of AIF1 binds the N-terminal RING finger domain of TRAF2, followed by ASK1-JNK signaling [623]. Thus, based on these findings, Luo *et al.* postulated that AIF1 may be directly involved in the IRE1α-TRAF2 complex and its activation of the ASK1-JNK signaling axis [321]. This hypothesis was validated in AIP1-KO mouse studies; AIP1-knockout mouse embryonic fibroblasts and vascular endothelial cells showed significant reductions in ER-stress induced ASK1-JNK activation that was rescued in AIP1 knock-in cells [321]. IRE1α has also been shown to directly interact with Bcl-2 family members Bax and Bak. Hetz *et al.* showed that Bax and Bak complex with the cytosolic domain of IRE1α and modulate IRE1α signaling [207]. Bax and Bak double knockout mice failed to signal through the IRE1α UPR branch following tunicamycin-induced ER stress; however, PERK signaling markers, e.g., eIF2α phosphorylation, responded normally [207]. This pro-activation role of Bak and Bax may be modulated by one of the few negative regulators of IRE1α activity, Bax inhibitor 1 (BI-1). BI-1 is an anti-

apoptotic protein that enhances cell survival following several intrinsic death stimuli [593]. Bailly-Maitre *et al.* were the first to suggest that BI-1 may down-regulate IRE1α and possibly ATF6 activity [21]. BI-1 deficient mice displayed increased XBP1s and enhanced JNK activity in the liver and kidney, while eIF2α phosphorylation remained normal under ER-stress conditions [21]. Lisbona *et al.* later showed that BI-1 directly interacts with the cytosolic domain of IRE1α, inhibiting its endoribonuclease activity [307]. Interestingly, BI-1 interacts with several members of the Bcl2 protein family e.g., Bcl2 and Bcl-X_L, even though it has no homology [593]. Members of the HSP family of proteins have also been shown to regulate IRE1α. For example, HSP90 interacts with the cytosolic domain of IRE1α, potentially protecting it from degradation by the proteasome [332]. HSP72 interaction with the cytosolic IRE1α domain has also recently been shown to enhance IRE1α endoribonuclease activity [175]. Taken together, these modes of IRE1α regulation with the exception of B1-1, largely promote or enhance IRE1α signaling.

2.2.5 ER stress-induced autophagy and apoptosis:

Ultimately, if UPR fails to restore ER homeostasis, cells initiate terminal programs such as autophagy or apoptosis. Several recent studies indicate that ER stress can trigger autophagy [386, 609, 36, 240, 218, 270, 146, 582]. Autophagy is an evolutionarily conserved cellular pathway, in which a cell recycles its macromolecules and organelles [295]. Autophagy is initiated by the formation of an autophagosome, composed of part of the cytosol or cellular organelles, encased in a double membrane [295, 276, 78]. Autophagosomes then bind endolysosomal vesicles, leading to the creation of the autolysosome. The autolysosome is

then degraded, completing the autophagy cycle. The key regulatory proteins involved in the nucleation and formation of the autophagosomal membrane are class III phospha-tidylinositol 3-kinase (PI3K), p150 myristylated protein kinase and Beclin-1 [295]. On the other hand, the ER-stress sensors PERK, IRE1 and cytosolic Ca^{2+} all act as effectors initiating autophagy in ER stressed cells [386, 218, 270, 146]. JNK1 mediated phosphorylation of Bcl2 at T69, S70 and S87 may also be important; phosphorylation at these sites leads to the dissociation of Bcl2 from Beclin-1, and the activation of autophagy [582]. Association of Beclin-1 with PI3K and other proteins promotes the localization of other autophagy related proteins to the preautophagosomal membrane [296]. Autophagy can be inhibited in eukaryotic cells by several factors, including the mammalian target of rapamycin complex 1 (mTORC1) protein complex [380, 348].

The autophagy program is a cellular reboot mechanism while apoptosis is a terminal death program. Apoptosis is characterized by nuclear and cytoplasmic condensation, blebbing of the plasma membrane and DNA fragmentation [247]. The dying cell eventually disintegrates into membrane-enclosed apoptotic bodies which are quickly destroyed by phagocytes or neighboring cells. A common biomarker of apoptosis is the activation of aspartate-specific proteases, collectively known as caspases [6]. Caspases rapidly dismantle cell cycle, cytoskeletal and organelle proteins by proteolytic cleavage. There are two pathways that result in caspase activation in response to apoptotic signals; the death-receptor and the stress mediated pathways. The death-receptor pathway is marked by ligand-mediated activation of death receptors on the plasma membrane. The alternative pathway for caspase activation is mediated by cellular stress e.g., ER stress. Caspases are activated from their zymogens (procaspases), in response to various death cues. First, the initiator caspases, caspase-8 and caspase-9, are

activated in response to death cues [360]. This is followed by the activation of executioner caspases, such as caspase-3, caspase-6 and caspase-7. Activated executioner caspases proteolytically process several substrates, facilitating cell death. They also activate initiator caspases, forming a positive feedback loop.

Activation of both the PERK and IRE1 pathways modulate stress-induced apoptosis through their regulation of Bcl2 expression and activity. Overall, stress induced apoptosis can occur through both mitochondrial-dependent and independent pathways. Stress signals cause oligomerization of pro-apoptotic proteins, such as Bax and Bak. These proteins are normally sequestered at the mitochondrial outer membrane by the survival protein Bcl2, under non-apoptotic conditions [581]. Once Bax and Bak oligomerize, they insert into the mitochondrial membrane and breach membrane integrity [370]. This results in a net efflux of cytochrome-c from the mitochondria to the cytosol and the initiation of the well-studied Apaf-1 mediated caspase-9 activation pathway. Recently, crosstalk between the ER and mitochondria during apoptosis, which might initiate mitochondrial apoptotic events, has also been explored [101]. This crosstalk may be facilitated via the release of Ca^{2+} ions from the ER into the cytosol, in response to ER stress. Mitochondrial uptake of Ca^{2+} initiates membrane fission and caspase activation via an uncertain mechanism. However, a few key drivers of ER Ca^{2+} release in response to stress signals have been identified [376]. Fas-mediated caspase-8 activation leads to cleavage of the membrane-bound ER protein, Bap31. Although Bap31 is membrane-bound, its caspase-recognition sequence is cytosolic, thus it is accessible by caspase-8. The truncated product, Bap20, is believed to promote the release of ER Ca^{2+} [108, 536, 101]. Interestingly, Bcl2 is also localized to the ER-membrane, but its role is uncertain [2]. Once in the mitochondria, Ca^{2+} modulates the permeability transition (PT)

pore (PTP) thereby increasing the efflux of cytochrome-c [99, 37]. Stress induced mitochondrial-independent apoptotic pathways are not well understood. Currently, caspase 12 has been suggested as a possible apoptotic mediator in mice [515, 608, 364]. However, caspase 12 is not expressed in human cells. Moreover, there is considerable debate about its role in stress-induced apoptotic cell-death [464].

2.3 The pathophysiology of ER-stress and aberrant UPR.

ER stress plays an important role in a spectrum of diseases ranging from neurodegeneration, cardiac diseases, cancer, diabetes to muscle degeneration (Table 2.3). Understanding ER-stress mechanisms in the context of these diseases presents a unique opportunity for drug discovery. Current therapeutic efforts have largely focused on amplifying the adaptive, pro-survival components of the UPR signal, for example, by inducing chaperone expression, which works to restore ER homeostasis. However, in the context of cancer, the opposite outcome is sought; in this case a number of strategies have focused on inhibiting proteasome function (Table 2.3.3).

2.3.1 Diabetes:

In the context of diseases such as diabetes, pancreatic β-cells depend on efficient UPR signaling to meet the demands for constantly varying levels of insulin synthesis. Type 1 diabetes is marked by excessive loss of pancreatic β-cells, while type 2 diabetes is marked by pancreatic β-cell dysfunction. The large biosyn-

Table 2.1: Relevance of ER stress in human disease states

Disease state	Role of ER-Stress and UPR	Key proteins	Reference
Alzheimer's disease (AD)	AD-induced Presenilin 1 regulates BiP.	Presenilin 1, PERK, eIF2α, BiP	[553, 242, 532, 352, 379]
Parkinson's disease (PD)	Suppressed ER-stress-induced apoptosis and aggregation of α-synuclein.	Parkin, α-synuclein	[92, 518, 224, 423, 461]
Amyotrophic lateral sclerosis (ALS)	Altered ERAD machinery and activation of ASK1.	SOD, ASK1 and Derlin-1	[378]
Bipolar Disorder	Current medications target induction of UPR.	XBP1	[239, 77, 238]
Stroke	CHOP and UPR mediates neuronal response in Ischemia.	PERK-eIF2α , ASK1, CHOP	[524, 281, 412, 516]
Heart Disease	Degeneration of cardiac myocytes, transaortic constriction. Myocardial infarction induces UPR.	IRE1, PERK-eIF2α, ASK1	[165, 388, 538, 496, 402, 567]
Atherosclerosis	Regulates inflammatory genes in vascular cells.	IRE1, JNK, TRAF2, XBP1	[152, 622, 289, 183]
Type 1 diabetes	Regulators in normal conditions and triggers β cell dysfunction and apoptosis in ER Stress.	PERK-eIF2α, JNK and Calcium	[475, 13, 135, 595, 227, 396, 397]
Type 2 diabetes	ER stress is a key aspect of type 2 diabetes.	XBP1, JNK	[399]
Cancer	Cytoprotective branches of UPR vital to survival and progression of tumors.	BiP, XBP1, ATF6, PERK	[498, 145, 230, 454, 40]
Autoimmune disease	Development of plasma cells and dendritic cells.	XBP1, GRP78, HLA-B27	[543, 43, 83]
Acute kidney injury	Upregulation of BiP, CHOP and down-regulation of Bcl-2 in primary glomerular diseases.	BiP, CHOP, Bcl2	[334, 22]

thetic load placed on the ER because of insulin production in response to food uptake (glucose) can overwhelm the folding capacity of the ER, resulting in ER stress. This leads to subsequent PERK activation and reduction of protein synthesis. In PERK -/- cells, protein synthesis is unresponsive to the stress and leads to accumulation of unfolded proteins (e.g. proinsulin) and ultimately cell death. It has been shown that PERK deficient mice are more prone to diabetes and progressive hyperglycemia [183]. In type II diabetes, ER stress leads to JNK-mediated phosphorylation of insulin receptor substrate 1 (IRS1) at S307 [399]. IRS1 is a substrate of Insulin receptor. Phosphorylation of IRS1 therefore inhibits insulin action. Nitric oxide (NO), is also a key player in β-cell death

in type-1 diabetes and vascular complications in type-2 diabetes. NO depletes ER Ca^{2+} leading to ER stress and ultimately apoptosis. Pancreatic β-cells have shown that NO-induced apoptosis is CHOP dependent [396]. Thus, ER stress is a critical feature of both type-I and type-II diabetes at the molecular, cellular and organismal level.

2.3.2 Role of UPR in hypoxia and cancers:

The ER not only acts as the center for maturation of proteins, but also as a critical node for oxygen sensing and signaling. In rapidly growing tumors, cells face stressors like hypoxia and nutrient deprivation both of which can lead to ER stress and ultimately UPR. For example, oxygen deprivation has recently been shown to be an initiator of UPR [269]. Interestingly, the connection between hypoxia and UPR is through hypoxia inducible factor (HIF) independent pathways. Hypoxia leads to PERK activation and the transient phosphorylation of eIF2α [268, 324, 265]. This effect occurs on a time scale of minutes in the case of anoxia, and is slower for moderate ($pO_2 \leq 1\%$) hypoxic conditions [265]. Because of their ability to transduce pro-apoptotic signals, the modulation of PERK and IRE1 activity has been explored as a potential anti-cancer strategy. Versipelostatin, a repressor of BiP expression, has been shown to produce anti-tumor activity in MKN-74 xenograft mouse models [410]. Enhanced apoptosis has also been observed in BiP-deficient fibrosarcoma cells, and XBP1- and PERK-deficient mouse fibroblasts [230, 454, 40]. Another related anti-cancer strategy is the induction of ER-stress using proteasome or protein trafficking inhibitors. Bortezomib, a selective inhibitor of the 26S proteasome, is used for the treatment of relapsed multiple myeloma, a plasma cell neoplasia [372]. Anti-

tumor activity was also observed with Bortezomib in preclinical pancreatic cancer studies using the L3.6pl cell-line [366]. Addition of Bortezomib to splenic and bone marrow plasma cells led to a 40-fold increase in CHOP expression, and a subsequent decrease in Bcl2 levels [372]. HIV type 1 (HIV-1) protease inhibitors (PI) like Nelfavir and Atazanavir are also currently being studied for the treatment of malignant gliomas [433]. Studies in glioblastoma cells suggested that Nelfavir or Atazanavir led to inhibition of proteasome activity and subsequently CHOP-induced cell death [433]. Thus, although no mechanistic link between proteasome inhibitors and UPR has been established, anecdotal evidence, such as CHOP induction, suggests these inhibitors are at least partially inducing ER-stress and subsequent UPR-mediated cell death. Recently, studies targeting intracellular trafficking of proteins have also been explored as an anti-cancer strategy. Brefeldin A treated chronic lymphocytic leukemia (CLL) cells showed induction of ER stress, caused by inhibition of protein trafficking, ultimately resulting in Golgi collapse and cell death [64]. Inhibition of protein trafficking from the ER can also lead to ER swelling and p53-independent apoptosis [64].

2.3.3 Neurodegenerative disorders:

Neurodegenerative diseases like Parkinsons or Alzheimers are often associated with the accumulation of unfolded or misfolded proteins. Incorrect protein folding alters neuronal connectivity, neuronal death and ERAD function [304, 506, 33, 243] (Table 1). Alzheimers disease (AD) is characterized by the progressive decline of cognitive function. This decline has been associated with the formation of protein aggregates involving several proteins including amy-

Table 2.2: Therapeutic interventions targeting ER stress.

Compound	Pathway	Function	Reference
Valporate (VPA), lithium	ER chaperones (BiP, GRP94, Calreticulin)	Remove ER Stress; ER chaperone induction in Neuro-protection in Bipolar Disorders.	[48, 210, 488]
BiP inducer X (BIX)	ER chaperones (BiP)	Remove ER Stress; Small-molecule inducer of BiP in Celebral ischemia; stroke resulting in reduced cell death.	[279]
Salubrinal	eIF2α, Bcl2	Phosphatase inhibitor; PERK-eIF2α pathway agonist in Parkinson's disease; anti-stroke; regulator of Bcl2.	[475, 49, 449, 281, 183, 13, 248]
p38 MAPK antagonists	p38 MAPK (CHOP activity)	Apoptosis up; p38 dependent CHOP phosphorylation regulator in stroke, diabetes.	[613, 574, 352, 412, 516]
Versipelostatin (VST)	BiP	Apoptosis up; Alters BiP production in glucose-deprived solid tumors.	[410]
Bortezomib, Nelfavir, Atazanavir	Proteasome, XBP1	Apoptosis up; Inhibition of HIV-1 proteasome activity. CHOP mediated cell death.	[288, 433, 160, 372]
Brefeldin A	CHOP, BiP	Apoptosis up; Inhibition of protein trafficking, ER swelling.	[64]

loid b-peptide (Ab) [483]. Mutations in the Ab precursor protein (APP) and Presenillins (PS1 and PS2), which mediate APP cleavage, are associated with the early onset of AD [587, 483, 242]. Mutant PS1 increased Ab production and ER stress mediated apoptosis in human neuroblastoma cells [564, 342, 282]. In these studies, the mutant PS1 suppressed IRE1 phosphorylation and ultimately down-regulated BiP expression [243]. However, there is contradictory evidence about the role of UPR in AD. For example, increased levels of BiP and PERK have been found in AD human brain specimens [216]. Thus, the central issue of whether UPR is initiated for neuro-protection or neuronal death, has yet to be determined.

2.3.4 Immune and Inflammatory disorders:

Unfolded or misfolded protein accumulation and UPR-induction is observed in autoimmune diseases like rheumatoid arthritis, inflammatory bowel diseases and multiple sclerosis. UPR and the inflammatory response are connected through several modalities including increased reactive oxygen species (ROS), increased calcium release from the ER, activation of JNK and the activation of NF-κB [628]. NF-κB plays a central role in inflammation [451]. In unstressed cells, NF-κB is inactive, sequestered by the inhibitor of NF-κB (IκB). UPR induction leads to phosphorylation and subsequent degradation of IκB and the expression of inflammatory genes [517, 451]. UPR also plays a key role in autoimmune diseases. CREBH, an ATF6 homolog, translocates to the golgi and undergoes cleavage by S1p and S2p proteases in the presence of ER stress. Proteolytic processing releases an N-terminal fragment that migrates to the nucleus and induces C-reactive protein and serum amyloid P-component expression [629]. Both of these genes are associated with the activation of the acute inflammatory response [629]. CREBH can also induce pro-inflammatory cytokines such as IL-6 and IL-1β [629].

2.3.5 The role of UPR in biologics production.

Biologics can be produced in a variety of hosts ranging from bacteria to mammalian cells [29]. Mammalian hosts, such as Chinese Hamster Ovary (CHO) cells, have been used for therapeutic protein production since as early as 1986 [592, 356]. Several current biologics are secreted, for example, the monoclonal antibodies (MAb) interferon-γ (IFNγ) or erythropoeitin (EPO) [278]. Thus, ER

processing and the unfolded protein response are critical to the production of these and many other therapeutic proteins. Interestingly, overexpression of ER chaperones like BiP and PDI typically have not significantly increased protein titers [47, 91, 260]. Instead, overexpression of other UPR proteins, such as X-Box binding protein 1 (XBP1), which is a key regulator of the secretory pathway, have had more pronounced effects on production [278, 539]. For example, overexpression of XBP1s increased by 2.5-fold MAb titers in CHO and NS0 cell lines which exhibit secretory bottlenecks [278]. Another strategy to enhance production is to modulate the translation block imposed by eIF2α phosphorylation. For example, GADD34 overexpression increased the de-phosphorylation of eIF2α, leading to an approximately 40% increase in the production of the glycoprotein Antithrombin III (ATIII) in CHO cells [394]. ATF4 overexpression has also been shown to increase ATIII production [387]. Thus, engineering mammalian hosts with optimal protein production capacity involves manipulating multiple unfolded protein response components, not just chaperones such as BiP.

2.4 Summary and Conclusions

Protein folding is strategically important to cellular function. Perturbations to the unique folding environment of the endoplasmic reticulum (ER) lead to the accumulation of unfolded or misfolded proteins. Higher eukaryotes have evolved a complex three-pronged response to correct aberrant folding, known as the unfolded protein response. The unfolded protein response acts to restore ER homeostasis. However, following prolonged or severe ER stress, cells initiate an alarm phase ultimately leading to apoptotic cell death. Malfunctions in the folding state of critical proteins have been linked with cancer, diabetes and

other diseases. Thus, the ability of UPR to regulate cell fate, has made it a primary area of study for pathophysiology and a potential therapeutic axis for the treatment of a diverse array of diseases.

CHAPTER 3

PHYSIOCHEMICAL MODELING AND ANALYSIS OF THE MAMMALIAN UNFOLDED PROTEIN RESPONSE

Authors - Chakrabarti A, Chen AW, Varner JD.

Submitted to - PLoS Comp. Biol.

3.1 Abstract

Cellular stressors routinely lead to malfunctions in the folding state of critical proteins, thereby playing a vital role in disease states like cancer, diabetes and cardiovascular ailments. Newly synthesized proteins must fold and assemble into unique three-dimensional structures, in order to become functionally active. Cells monitor protein folding by an inbuilt quality-control system involving both the Endoplasmic Reticulum (ER) and the Golgi apparatus. Incorrectly folded proteins are tagged for degradation via ER associated degradation (ERAD) or sent back through a refolding cycle. However, accumulation of incorrectly folded proteins can trigger a cascade of events, termed the Unfolded Protein Response (UPR). UPR leads to re-establishing the cellular homeostasis. In this study, we developed a family of mechanistic models of eukaryotic UPR, which was composed of a system of ordinary differential equations. The objective of this study was to assemble a series of molecular modules describing different aspects of UPR and subsequently analyze these models for fragility and robustness. Kinetic parameters for these models were estimated by comparing simulations with experimental data. Using POETs and a cross-validation scheme, we developed an ensemble of models, consistent with literature data.

Model analysis highlighted the presence of a sequential order for firing of the ER stress transducers in UPR. The counteracting effects of the ER stress transducers converge at feedback regulation of the molecular chaperone BiP. The three main regulators of feedback, ATF4, cleaved ATF6 and XBP1s share the load ensuring the phased response of UPR. However these regulators add on to the overall fragility of the system and allow scope for manipulation of UPR as seen by sensitivity/robustness analysis. Downstream effects of UPR include a balance between cell survival and cell death based on the magnitude of the stress. Model analysis suggested that the cell-death axis was relatively robust owing to redundant routes e.g., APAF-1 dependent and APAF-1 independent routes of apoptosis. However the cell-survival axis was relatively susceptible to perturbations owing to multiple levels of regulation, thereby highlighting importance of key proteins like CHOP.

3.2 Introduction

Protein folding is strategically important to cellular function in all organisms. In eukaryotes, secreted, membrane-bound and organelle-targeted proteins are typically processed and folded in the endoplasmic reticulum (ER) [363, 456, 245]. Intracellular perturbations caused by a variety of stressors disturb the specialized environment of the ER leading to the accumulation of misfolded or unfolded proteins [117, 134]. Shifts in folding capacity have been associated with diseases such as cancer, diabetes and cardiovascular disorders [117]. Physiological processes such as aging can also influence protein folding [363]. Normally, cells ensure proper protein folding using a combination of molecular chaperones, foldases and lectins [363]. However, when proper folding can not be re-

stored, unfolded or misfolded proteins are targeted to ER Associated Degradation (ERAD) pathways for processing [245]. If unfolded or misfolded proteins continue to accumulate, eukaryotes induce the unfolded protein response (UPR). In mammalian cells, UPR is a complex signaling program mediated by three ER transmembrane receptors: activating transcription factor 6 (ATF6), inositol requiring kinase 1 (IRE1) and double-stranded RNA-activated protein kinase (PKR)-like endoplasmic reticulum kinase (PERK). UPR performs three functions, adaptation, alarm and apoptosis. During adaptation, the UPR tries to reestablish folding homeostasis by inducing the expression of chaperones that enhance protein folding. Simultaneously, translation is globally attenuated to reduce the ER folding load while the degradation of unfolded proteins is increased. If these steps fail, the UPR induces a cellular alarm and apoptosis program. The alarm phase involves several signal transduction events, ultimately leading to the removal of the translational block and the down-regulation of the expression and activity of pro-survival factors such as the B-cell lymphoma 2 (Bcl2) protein. After the alarm phase, cells can undergo apoptosis, although ER stress can also initiate autophagy [386, 609, 36, 240, 218, 270, 146]. Thus, ER folding homeostasis strongly influences mammalian physiology [134].

In this study we used physiochemical computer modeling as a tool to study the mammalian unfolded protein response. Physiochemical models which describe system-level responses can prioritize experimental directions, generate testable hypothesis and perhaps identify and validate potential therapeutic targets [256]. For mammalian UPR, physiochemical modeling might tell us the critical components of each branch, and ultimately what must be manipulated to get a desired network response, for example, enhanced death or survival. However, there are critical limitations with physiochemical models. Foremost

amongst these is the large number of unknown model parameters and uncertainty in the model structure. It is typically impossible to uniquely identify model parameters, even with extensive training data and perfect models [148]. Alternatively, ensemble approaches, which use uncertain model families, have emerged as a promising strategy in systems biology and other fields like climate prediction [26, 280, 54, 401, 505]. Their central value has been the ability to quantify simulation uncertainty and to experimentally constrain model predictions. For example, Gutenkunst *et al.* showed that predictions were possible using signal transduction model ensembles, despite sometimes only order of magnitude parameter estimates [176]. More recently, Luan *et al.* predicted patient response to therapeutic intervention using an ensemble of human coagulation models where approximately 40% of the model parameters had a coefficient of variation (CV) greater than 0.5 [319]. Model ensembles have also been used with parameter dependent analysis techniques, e.g., sensitivity analysis to robustly estimate important signaling network features. For example, Tasseff *et al.* characterized emergent behavior between androgen and growth factor signaling in prostate cancer cell-lines using an ensemble of models and sensitivity analysis [526]. Thus, despite uncertainty, ensembles of physiochemical models can be used to understand qualitative properties of complex biochemical networks.

We developed a population of physiochemical models describing the adaptation, alarm and apoptosis phases of mammalian UPR. Traditionally, it has been hypothesized that there is a sequential order for firing of the ER stress transducers in UPR. The PERK and ATF6 branches are thought to be activated before IRE1 [515] and largely promote ER adaptation to misfolding, while IRE1 is the last one to fire as it has a dual role, transmitting both survival and pro-apoptotic signals. However our modeling analysis suggests that these three

branches fire simultaneously with varying rates and the state of the cell in terms of adaptation, alarm or apoptosis is a result of counteracting effects of these three prongs of UPR signaling. The counteracting effects is further substantiated by simulated knockout studies, where in knockout of one ER stress transducer lead to enhancement of the other branches of UPR, e.g, if we want more protein production even with ER stress, PERK KO might suggest as a viable route as that shall lead to no translational block. However the IRE1 branch shall amplify and so will the ATF6 branch and ultimately still lead to death (less compared to WT). So a viable strategy could be to regulate CHOP, as it shall lead to reduced death by regulation of Bcl2. One common playground for the three ER stress transducers is the feedback regulation of BiP expression. Model analysis substantiated the significance of BiP feedback in the overall robustness of the system. Removal of any one branch of BiP feedback lead to increased sensitivity of the other branches. So there was load/responsibility sharing within the system. Interestingly, removal of all nodes of BiP feedback increased the overall robustness of the system. So while BiP feedback is crucial in terms of the cell to adapt to small perturbations, at the same time this makes the system more fragile and susceptible to manipulations. Model analysis further highlighted the presence of redundant routes of regulation of apoptosis such as APAF-1 dependent/independent strategies. While manipulation of pro-death axis of the system (cleaved caspases) was relatively robust to perturbations, manipulation of the pro-survival axis (Bcl2) was relatively feasible via CHOP dependent/independent strategies. Overall UPR was seen to be robust to perturbations (simulated knockout/overexpression studies), thereby highlighting the redundancy and crosstalk within the three branches of UPR. However, BiP regulation at the transcriptional level via intermediates of the ER-stress transducer

signaling cascades (e.g. ATF4, XBP1s) was seen to be the key in the regulation of UPR. The EUPR model is available in SBML in the supplemental materials.

3.3 Results

3.3.1 Formulation of the UPR network architecture.

The UPR network described the ER folding cycle, ER-associated degradation (ERAD), ER-stress transducer (PERK, IRE1 and ATF6) signaling cascades and stress-induced caspase activation (Fig. 3.1). The network consisted of 636 protein or mRNA species interconnected by 1090 interactions (Fig. 3.1 Inset). Connectivity was formulated from a comprehensive review of the primary literature [363, 456, 245, 117, 134, 481, 167, 345, 32, 208, 554, 515], and from on-line databases; String-8 [233], NetworKIN [305] and TRANSFAC. Model connectivity was not specific to a single cell-line. Rather, it was a canonical representation of the pathways involved in monitoring and controlling the folding capacity of a generic well-mixed ER compartment. UPR induction was modeled as the release of BiP from the ER stress transducers, PERK, IRE1α and ATF6 leading initially to adaptation of the folding cycle and then subsequently to alarm and apoptosis. The adaption phase of UPR was marked by general translation attenuation, selective transcriptional programs for key species like bZIP transcription factor ATF4 [318], cellular inhibitor of apoptosis (cIAP) [179], molecular chaperones e.g., BiP [185] and enhanced clearance of accumulated proteins via ERAD. The alarm and apoptosis phases were mediated by the induction of CHOP [455], regulation of Bcl2, Bcl2-antagonist of cell death (BAD) [571] and (TNF) receptor

associated factor 2 (TRAF2) [294, 430, 596, 515] activation (Fig. 3.1). Model connectivity is available in SBML format from the supplemental materials.

3.3.2 Estimating a population of canonical UPR models using POETs.

Despite a significant identifiability challenge, the multiobjective POETs algorithm generated a predictive UPR model population. The three phases of UPR were modeled using mass action kinetics within an ordinary differential equation (ODE) framework. While ODEs and mass-action kinetics are common methods of modeling biological pathways [479, 74, 203], this modeling strategy resulted in a large number of unknown model parameters. These parameters were not uniquely identifiable (data not shown). Instead, we estimated an experimentally constrained population of parameters using multiobjective optimization. A population of the 1726 unknown model parameters (1090 kinetic parameters and 636 initial conditions) was estimated from 33 dynamic and steady state data-sets taken from literature (Table 3.1). The residual between model simulations and each of the experimental constraints was simultaneously minimized using the multiobjective POETs algorithm [504]. A leave-eight-out cross-validation strategy was used to independently estimate the training and prediction error over the 33 data sets; we estimated four different model families, where eight of the 33 objectives were reserved for validation and 25 were used model training each family. Thus, each model family was trained and validated on different experimental data. Starting from an initial best-fit initial parameter set (nominal set), more than 25,000 probable models were estimated

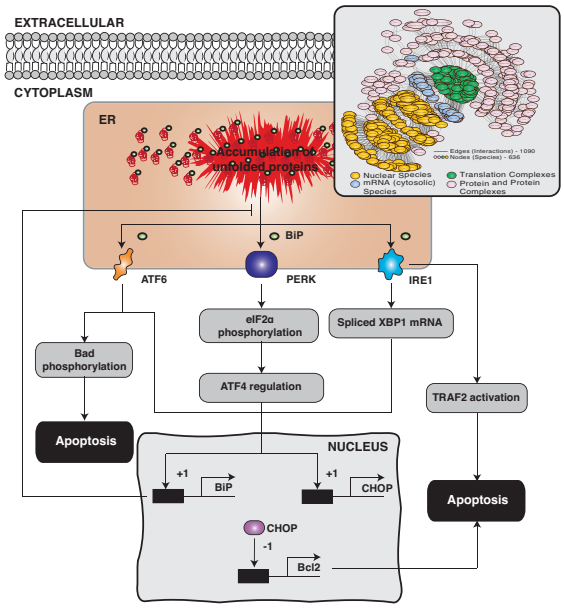


Figure 3.1: An array of cellular stressors can perturb the folding environment in the endoplasmic reticulum (ER) leading to unfolded or misfolded protein. In response to the folding imbalance, cells initiate the cytoprotective unfolded protein response (UPR). The problem of unfolded or misfolded proteins in the ER is addressed by increasing the folding capacity through the up-regulation of the expression of chaperone proteins, attenuating translation by regulating eIF2α, and promoting the degradation of misfolded proteins through ER-associated degradation (ERAD). If UPR is unable to restore the folding balance, ER stress will eventually lead to apoptotic cell-death. The three signal transduction pathways mediating the unfolded protein response in higher eukaryotes. First, the PRKR-like ER kinase (PERK) pathway is initiated after BiP dissociation from PERK. While PERK transduces both pro- and anti-apoptotic signals, its main function is translation attenuation through the phosphorylation of eIF2α. Next, the activating transcription factor 6 (ATF6) pathway is activated following BiP dissociation. ATF6 induces the expression of chaperones e.g., BiP as well as apoptosis effectors such as CHOP. Lastly, the inositol-requiring kinase 1 (IRE1) pathway is activated following BiP dissociation from IRE1. Activated IRE1 has both an endoribonuclease and a serine-threonine kinase activity that drive can pro-apoptotic signals. Inset: The UPR network consisted of 636 protein or mRNA species interconnected by 1090 interactions.

by POETs from which we selected N = 100 models (25 from each training family) with a Pareto rank of one or less (from approximately 1200 possible choices) for further study. For each objective, the nominal, training (75 models) and prediction (25 models) error was calculated for each objective (Table 3.1). Models used for prediction error calculations for a particular objective were *not* trained on that objective. The prediction likelihood was statistically significantly better for 31 of the 33 objective functions at a 95% confidence level, compared with random parameter sets generated from the nominal set (Table 3.1). Thus, POETs generated model families that predicted approximately 94% of the objective functions with a significantly higher likelihood than a random control. However, the specific value of any given parameter was likely not well described. The coefficient of variation (CV) for the model parameters ranged from 0.5 - 1.6, where approximately 65% of the parameters were constrained with a CV $\leq$ 1.0 (Fig. 3.2). The most constrained parameters involved a wide-array of functions e.g., regulation of PERK, eIF2α, ATF4, Calcineurin, BiP, CHOP and ATF6 signaling. However, the least constrained parameters involved JNK and apoptosis interactions.

POETs identified Pareto fronts between several objectives, e.g., O13×O14, O25×O29, O11×O29 and O27×O2 in the training data (Fig. 3.3). Strong Pareto fronts suggested an inability to simultaneously model different aspects of the training data. However, fronts could also result from experimental artifacts, e.g., variation between cell-lines, time-scale differences or from functional relationships in the data. Globally, adaptation and alarm phase training constraints conflicted with those involving apoptosis. For example, objectives involving caspase-7 or caspase-9 activity conflicted with phosphorylated eIF2α levels. Phosphorylation of eIF2α by activated PERK attenuates translation, which de-

Table 3.1: Objective function list along with species, cell-type, nominal error, training error, prediction error, random error with a randomly generated parameter set and the corresponding literature reference.

Obj#	species	cell type	nominal	training	prediction	random	source
O1	ATF6 free	HEK293	0.25	0.25 ± 0.04	0.26 ± 0.08	0.49 ± 0.05	[301]
O2	peIF2α	HEK293	0.14	0.23 ± 0.08	0.25 ± 0.11	0.51 ± 0.11	[301]
O3	BiP mRNA	HEK293	0.24	0.43 ± 0.17	0.49 ± 0.15	0.72 ± 0.13	[301]
O4	CHOP mRNA	HEK293	0.55	0.56 ± 0.07	0.57 ± 0.06	0.68 ± 0.20	[301]
O5	Cleaved PARP	HEK293	0.24	0.23 ± 0.14	0.23 ± 0.17	0.50 ± 0.19	[302]
O6	pIRE1α	AR42J	0.37	0.36 ± 0.05	0.35 ± 0.07	0.63 ± 0.09	[38]
O7	PERK	AR42J	0.19	0.36 ± 0.27	0.38 ± 0.29	0.46 ± 0.20	[38]
O8	pPERK	AR42J	0.10	0.14 ± 0.09	0.15 ± 0.15	0.28 ± 0.08	[38]
O9	BIP mRNA	MEF	0.19	0.30 ± 0.14	0.26 ± 0.07	0.66 ± 0.24	[597]
O10	BIP Protein	P19 EC	0.42	0.34 ± 0.15	0.35 ± 0.11	0.68 ± 0.15	[236]
O11	CHOP Protein	P19 EC	0.24	0.42 ± 0.20	0.38 ± 0.18	0.59 ± 0.20	[236]
O12	sXBP1 protein	MEF	0.34	0.29 ± 0.08	0.30 ± 0.06	0.40 ± 0.11	[60]
O13	CHOP Protein	MEF	0.46	0.43 ± 0.12	0.45 ± 0.13	0.56 ± 0.18	[60]
O14	XBP1 mRNA	MEF	0.57	0.48 ± 0.13	0.49 ± 0.13	0.47 ± 0.05	[60]
O15	peIF2α	MEF	0.38	0.37 ± 0.08	0.38 ± 0.03	0.44 ± 0.11	[598]
O16	ATF4 nuclear	MEF	0.63	0.67 ± 0.06	0.67 ± 0.07	0.68 ± 0.16	[598]
O17	CHOP Protein	MEF	0.52	0.51 ± 0.12	0.53 ± 0.08	0.64 ± 0.19	[598]
O18	BIP Protein	MEF	0.18	0.34 ± 0.18	0.34 ± 0.23	0.55 ± 0.16	[598]
O19	Spliced XBP1	MEF	0.49	0.49 ± 0.04	0.48 ± 0.04	0.50 ± 0.10	[291]
O20	Cleaved ATF6	MEF	0.45	0.44 ± 0.06	0.44 ± 0.05	0.49 ± 0.11	[291]
O21	cIAP Protein	NIH 3T3	0.36	0.25 ± 0.15	0.30 ± 0.10	0.52 ± 0.15	[179]
O22	Caspase 12 Protein	Transfected HEK 293T	0.44	0.57 ± 0.11	0.52 ± 0.11	0.75 ± 0.11	[438]
O23	Phospho eIF2α	Liver of Mice	0.50	0.51 ± 0.07	0.51 ± 0.04	0.50 ± 0.12	[206]
O24	Procaspase 12	SAK2	0.28	0.46 ± 0.28	0.38 ± 0.26	0.78 ± 0.14	[440]
O25	Caspase 12	SAK2	0.12	0.22 ± 0.12	0.25 ± 0.13	0.73 ± 0.07	[440]
O26	Procaspase 9	SAK2	0.14	0.28 ± 0.22	0.33 ± 0.26	0.78 ± 0.13	[440]
O27	Caspase 9	SAK2	0.03	0.19 ± 0.16	0.24 ± 0.18	0.68 ± 0.14	[440]
O28	Procaspase 7	SAK2	0.31	0.44 ± 0.24	0.43 ± 0.23	0.74 ± 0.14	[440]
O29	Caspase 7	SAK2	0.64	0.61 ± 0.07	0.58 ± 0.11	0.75 ± 0.07	[440]
O30	Procaspase 3	SAK2	0.34	0.46 ± 0.19	0.46 ± 0.21	0.82 ± 0.13	[440]
O31	Caspase 3	SAK2	0.34	0.39 ± 0.09	0.42 ± 0.13	0.74 ± 0.10	[440]
O32	PARP Protein	SAK2	0.37	0.51 ± 0.17	0.51 ± 0.14	0.85 ± 0.13	[440]
O33	Cleaved PARP	SAK2	0.11	0.21 ± 0.17	0.31 ± 0.21	0.67 ± 0.15	[440]

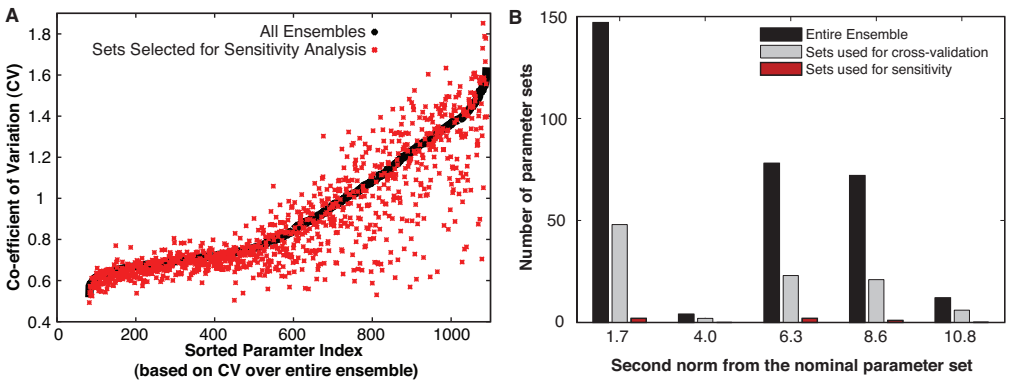


Figure 3.2: POETs generated an ensemble of models that predicted approximately 94% of the objective functions with a significantly higher likelihood than a random control. (A) The coefficient of variation (CV) for the model parameters ranged from 0.5 - 1.6, where approximately 65% of the parameters were constrained with a CV $\leq$ 1.0 (black dots). (B) We selected five parameter sets (red dots in A) for further analysis based on CV and distance from the nominal parameter set (based on second norm).

creases the ER folding load. Thus, eIF2α phosphorylation is a key early adaptive event in UPR. On the other hand, caspase-9 is a stress-induced death marker activated only after UPR has failed to restore folding homeostasis. Conflicts between these early and late phase markers suggested the UPR time-scale was perhaps cell-line or perturbation dependent. Similarly, negative feedback may lead to conflicting objectives. For example, XBP1 mRNA measurements (O14) conflicted with CHOP protein measurements (O13), even though these data-sets were taken from the same study and were collected in the same cell-line. XBP1 splicing increased BiP levels, which in turn reduced CHOP protein levels, hence the trade-off. Lastly, in addition to fronts, we also observed strong correlation between objectives. For example, models that performed well for the CHOP protein (O11), also performed well against Procaspase-12 (O22) measurements, even though these were not in the same cell-line or from the same study. Both CHOP and Procaspase-12 are downstream of IRE1/TRAF2/JNK signaling cas-

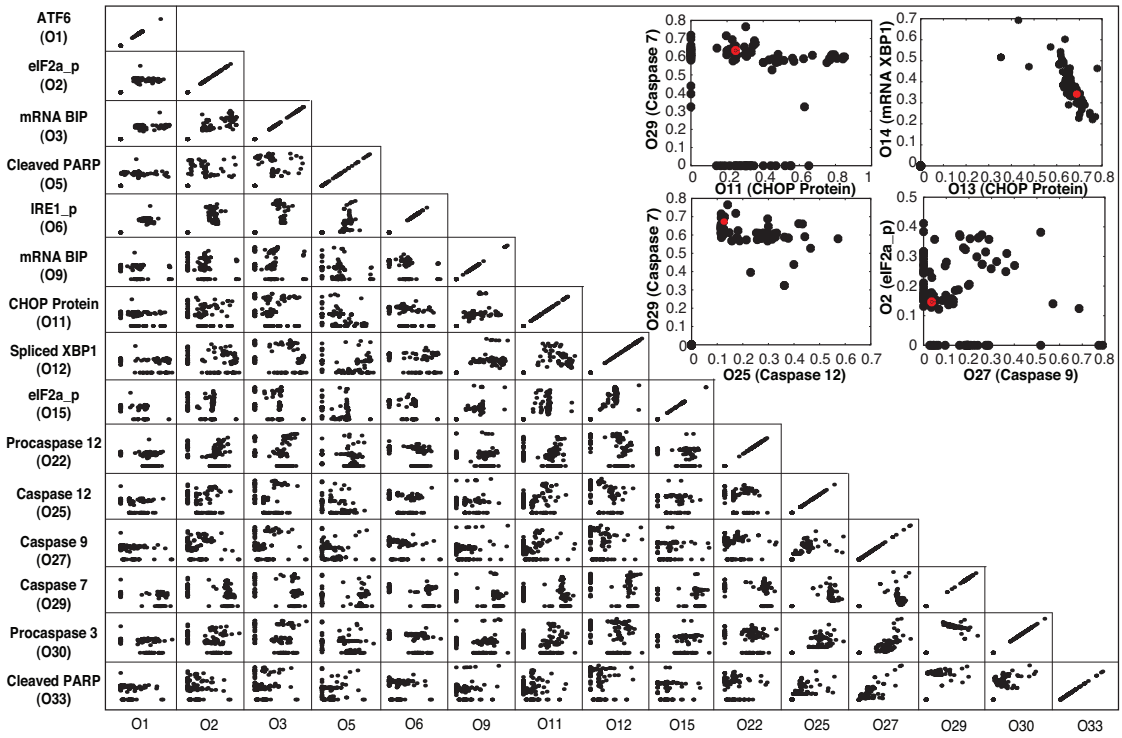


Figure 3.3: Objective function plot for selected training constrains (O1,O2,...O33) for the UPR model population generated using POETs. Points denote separate models in the population. Several objectives exhibit clear Pareto fronts, e.g., O29 × O25. This suggests an inability to model both training constraints simultaneously or conflicts in the training data.

cade, so these errors were directly correlated.

3.3.3 The population of UPR models recapitulated adaptation, alarm and apoptotic events across multiple cell-lines and timescales.

We trained the UPR ensemble using data from UPR initiation events (e.g., PERK activation) and the downstream activation of proteins involved in apoptosis (BiP, caspase-3 and caspase-7). ER stress induced by exposure to thapsigargin (Tg), a non-competitive inhibitor of SERCA Ca^{2+} transporters, leads to the dissociation of BiP from the ER-stress transducers. For theses initial proof-of-concept simulations, we assumed that the action of Tg and other stress-inducing agents such as dithiothreitol (DTT) was identical, i.e., induction of BiP dissociation. Within 20 min after Tg exposure, PERK was activated and transmitting adaptation signals downstream. The population of UPR models recapitulated the timescale of PERK phosphorylation (Fig. 3.4A) as well as its downstream signaling activity, for example, the phosphorylation of eIF2α (Fig. 3.4H). The nuclear fraction of ATF4 increased from approximately zero (untreated cells) to a maximum value 4 hrs after Tg exposure. While the model ensemble generally predicted the correct trend, there was significant error in the early time points for ATF4 (Fig. 3.4G). The phosphorylation of eIF2α by PERK is required for ATF4 activation. Interestingly, when we compared model simulations of p-eIF2α levels following Tg (1μM) exposure in MEFs with measurements (O15 [598]), the model correctly captured the appropriate behavior. To test the functioning of the ATF6 branch of UPR model, we compared simulations with measurements of cleaved ATF6 in tunicamycin-treated MEFs [291]. ER stress leads to the release of the BiP from ATF6. Cleaved ATF6 is then translocated to the nucleus where it up-regulates gene expression [191, 178]. Simulations of cleaved ATF6

levels following UPR initiation were consistent with measurements (Fig. 3.4C). The formation of p-PERK or p-IRE1 initiates a complex series of events that operate on both short and long time-scales. Signals from the ER stress transducers converge downstream to regulate BiP transcription [330, 439, 263, 612]. Model ensemble recapitulated correct trends of BiP mRNA with maximum levels ~ 8 hrs, similar to what was seen in experiments done on HEK293 cells (Fig. 3.4D) [301]. One of the long-term outcomes of PERK/IRE1 activation is apoptotic cell death. In the proof-of-concept ER-model the link between UPR and apoptosis occurred through the action of eIF2α, the dual role of the ATF4 transcription factor and caspase-12 activation by IRE1-TRAF2 signaling axis. We constrained model parameters associated with the activation of the cell-death program using measurements of pro/caspase-7 levels, pro/caspase-9 levels, pro/caspase-3 levels, pro/caspase-12 levels and PARP cleavage mediated by executioner caspases following treatment with 0.5μM Tg [440]. These experiments were performed in Sak2 cells that lack Apaf-1 protein expression [440]. Thus, these data allowed us to include a non-Apaf-1 mediated stress-induced caspase activation pathway into the model. The population of models recapitulated caspase-3 (Fig. 3.4K), caspase-9 (Fig. 3.4J) as well as cleaved Parp levels (Fig. 3.4L) following exposure to ER stress-inducers. Interestingly, while PERK activation occurred on the timescale of minutes, initiator and executioner caspase activation occurred over 36 hrs. Thus, the population of UPR models captured complex signaling events occurring across multiple timescales.

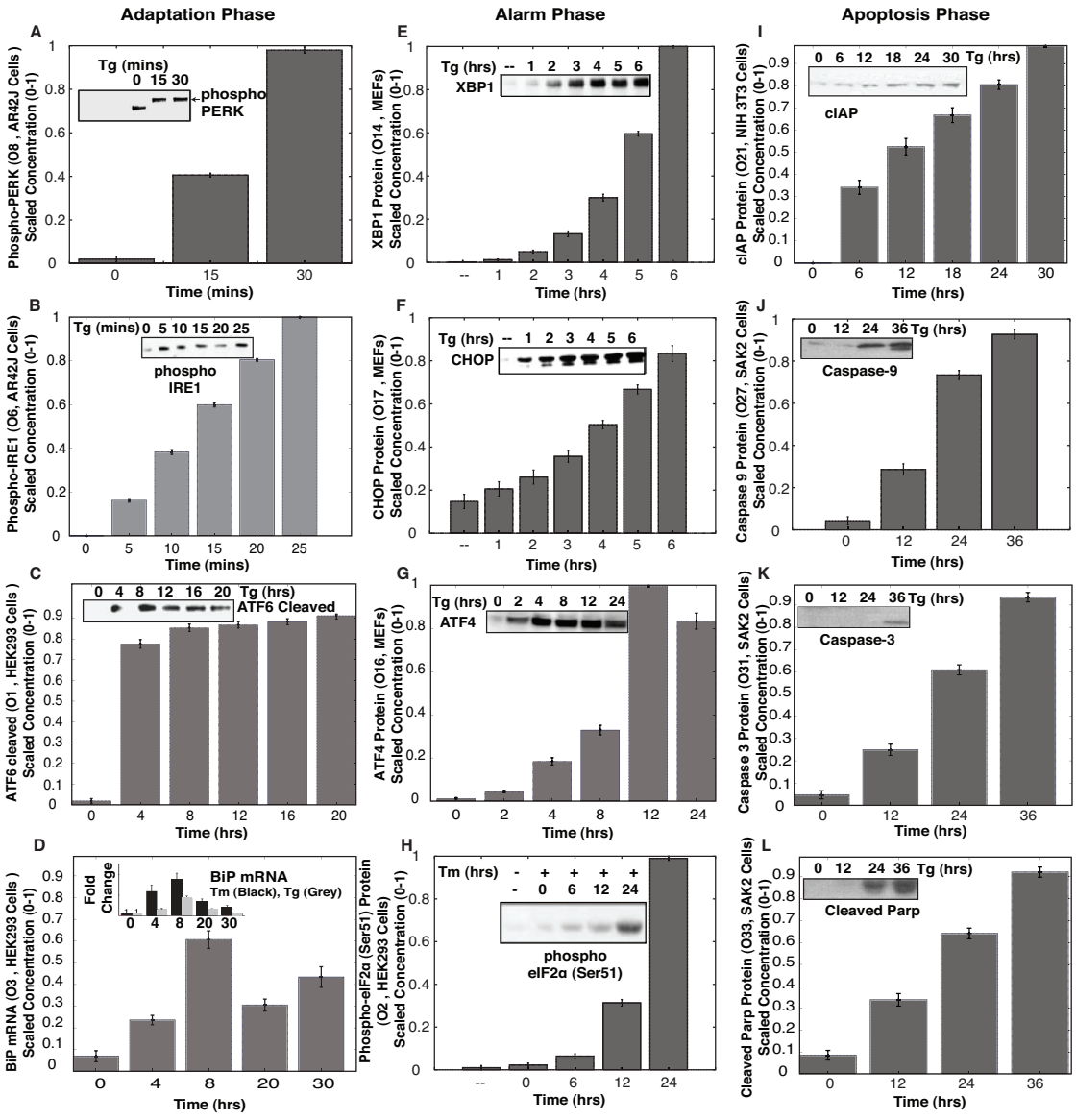


Figure 3.4: Simulations versus experimental data for selected objective functions following exposure to the ER-stress inducers Thapsigargin (Tg or Thaps) or Tunicamycin (TM). The first-column (A - D) denotes adaptation components, the second column (E - H) denotes alarm phase components, while the third column (I - L) denotes apoptosis phase components. Bars denote the scaled mean concentration computed over the ensemble, while the error bars describe one standard error.

3.3.4 Signal flow analysis of UPR highlighted modes of crosstalk and redundancy in BiP and Bcl2 regulation.

Traditionally, it has been hypothesized that there is a sequential order for firing of the ER stress transducers in UPR. The PERK and ATF6 branches are thought to be activated before IRE1 [515] and largely promote ER adaptation to misfolding, while IRE1 is the last one to fire as it has a dual role, transmitting both survival and pro-apoptotic signals. However our modeling analysis suggests that these three branches fire simultaneously with varying rates and the state of the cell in terms of adaptation, alarm or apoptosis is a result of counteracting effects of these three prongs of UPR signaling (Fig. 3.5 and Fig. 3.6).

The counteracting effects is further substantiated in simulated knockout studies, where in knockout of one ER stress transducer leads to enhancement of the other branches of UPR (Fig. 3.6A-C). UPR induction in the model was controlled by manipulation of the generation rate of unfolded or misfolded protein (qP) in the ER compartment. Upon UPR induction, initially ($\leq$ 1hr) the response is damped marking the adaptation phase of UPR. This is followed by an increase in the activity of the IRE1α, PERK and ATF6 cascades at $\sim$ 1 hr marking the onset of the alarm phase, which leads to a final steady state $\sim$ 8-10 hrs marking the onset of the commitment or apoptosis phase of UPR (Fig. 3.5). This time frame was consistent with the maximum levels of BiP mRNA upon UPR induction [301]. This analysis was substantiated further by looking at the fluxes at different phases of UPR induction. As compared to P1 (No-UPR Steady State Fig. 3.5), we saw that early on at 1 hr after unfolded protein dose there was a marked increase in ATF4 and CHOP regulation, ATF6 signaling along with unfolded protein sensing and degradation. These are hallmarks of the adaptation-

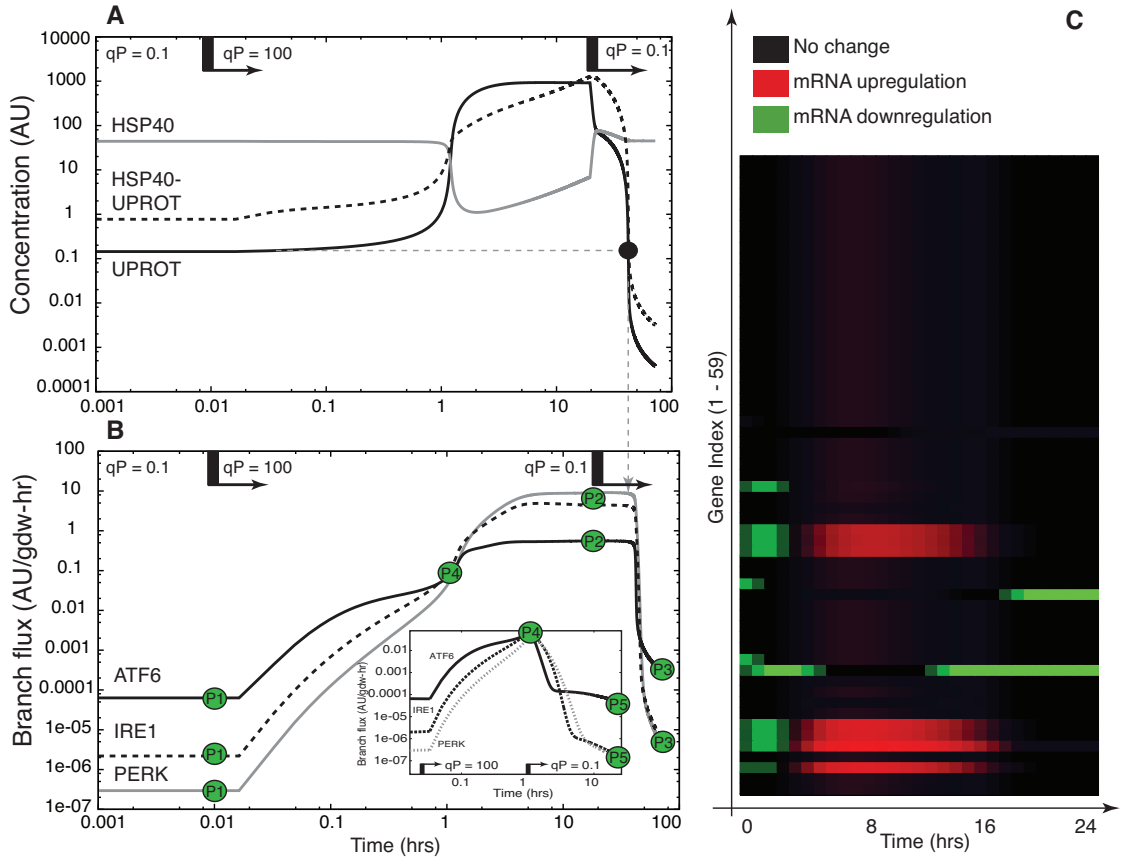


Figure 3.5: Proof of concept simulation unfolded protein response activation. **A:**UPR induction was controlled by manipulating the generation rate of unfolded or misfolded protein (qP) in the ER compartment. A step-change in qP from qP = 0.1 to qP = 100 was issued at approximately t = 0.1 hrs and then adjusted back to qP = 0.1 at t = 20 hrs. **B:**Flux through the PERK, ATF6 and IRE1 stress sensing branches as a function of time following a step change in misfolded protein generation. **C:**Simulated expression profile for the 59 genes in the model. The symbol UPROT denotes the level of unfolded protein.

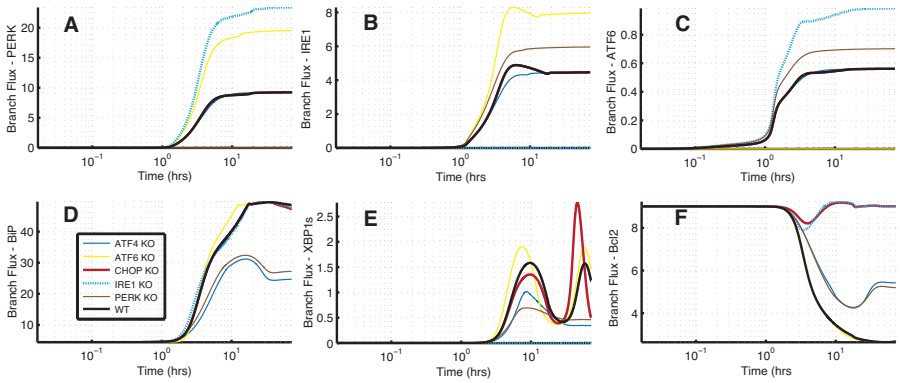


Figure 3.6: Signal flow analysis using simulated knockout (KO) of key proteins on the UPR system: Simulation results suggest that the three branches in UPR fire simultaneously with varying rates and the state of the cell in terms of adaptation, alarm or apoptosis is a result of counteracting effects of these three prongs of UPR signaling. (A-C) The counteracting effects is seen when knockout of one ER stress transducer leads to enhancement of the other branches of UPR. (D) ATF4, cleaved ATF6 and XBP1s act as integrators of the signals coming from all the three branches of UPR and furthermore leads to regulation of BiP, thereby leading to a negative feedback or control of UPR signal. PERK and ATF4 KO studies revealed a slower and lower amount of BiP production (~ 50%) as compared to WT. However, ATF6 or IRE1 KO did not affect BiP regulation as compared to WT. (E) Regulation of BiP was the critical regulator of spliced XBP1 (XBP1s), which in turn acts as a key marker of progression through different stages of UPR. (F) PERK and ATF4 KO lead to delay in the onset of apoptosis (marked by slower and lower reduction of Bcl2 levels. This effect could be attributed to the lack of CHOP mediated branch of Bcl2 regulation. On the other hand, IRE1 and CHOP KO leads to drastic reduction in apoptosis (marked by little or no change of Bcl2 levels). CHOP KO, implicated the importance of CHOP in the down-regulation of Bcl2. IRE1 KO implicated the critical role of IRE1-TRAF2 mediated route of apoptosis. Overall flux analysis highlighted the extensive amount of crosstalk within the three branches of the UPR network.

alarm phase of the UPR response (Fig. 3.7D). The apoptosis phase (≥ 8-10 hrs P2), was marked by increased BiP regulation, enhanced ATF4 transcriptional activity, increased mitochondrial membrane permeability and increased apoptotic fluxes. So the cell has committed itself to apoptosis mediated cell death. We compared the ability of the system to recuperate upon reversal of unfolded protein loads from the alarm phase (P4, Fig. 3.5) and the apoptosis phase (P2, Fig. 3.5). If we reduced the protein load after the cell was committed to apoptosis (as in P3), we found that the cell continued to function similar to the UPR state even upon protein load reduction after 25 hrs (Fig. 3.7C). Certain modules were seen to reduce like IRE1-TRAF2 signaling, ASK1 activation. However not much difference was seen in terms of apoptotic fluxes, denoting the commitment of the cell to death and to a point of no return. On the contrary if we reduced the load of proteins in the adaptation-alarm phase (P4), we saw that the cell could recuperate using its ERAD machinery and the regulation of BiP (Fig. 3.7E-F).

We further investigated the effect of simulated knockout (KO) and overexpression (OX) of key proteins on the UPR system. Interestingly, PERK and ATF4 KO studies revealed a slower and lower amount of BiP production ($\sim$ 50%) as compared to WT. However, ATF6 or IRE1 KO did not affect BiP regulation as compared to WT. This highlighted the dominant role of ATF4 in regulation of BiP, which is consistent with experimental evidence [324]. Regulation of BiP was the critical regulator of spliced XBP1 (XBP1s), which in turn acts as a key marker of progression through different stages of UPR (Fig. 3.6E). ATF4, cleaved ATF6 and XBP1s act as integrators of the signals coming from all the three branches of UPR and furthermore leads to regulation of BiP, thereby leading to a negative feedback or control of UPR signal. Another interesting note was the regulation of pro-apoptosis phenotype via regulation of Bcl2. PERK and ATF4 KO lead to

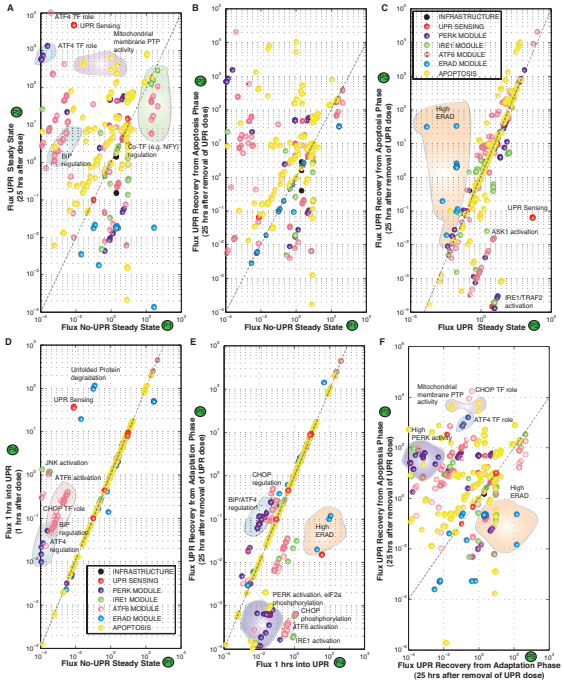


Figure 3.7: Cross plot of the fluxes at P1-P5 as denoted in Fig. 3.5: We tried to see is how the system behaves and how the system can recuperate from UPR dose when it is in the adaptation phase as compared to the apoptosis phase. (D) As compared to P1 (No-UPR Steady State), we see that early on at 1 hr after UPR dose there is a marked increase in ATF4 and CHOP regulation, ATF6 signaling along with unfolded protein sensing and degradation. These are hallmarks of the adaptation-alarm phase of the UPR response. (A) If we continue with the dose of UPR till around 25 hrs, we see the fluxes reach a steady state. This state is marked by increased BiP regulation, enhanced ATF4 transcriptional activity, increased mitochondrial membrane permeability and increased apoptotic fluxes. This state is similar to the Apoptotic phase of UPR, where in the cell has committed itself to apoptosis mediated cell death. (B) and (C) If we reduce the UPR load after the cell has committed to apoptosis (as in P3), we find that the cell continues to function similar to the UPR state even upon UPR load reduction after 25 hrs. There are certain aspects which are seen to reduce like IRE1-TRAF2 signaling, ASK1 activation. However not much difference is seen in terms of apoptotic fluxes, denoting the cell has committed itself to death and is in a point of no return. (E)-(F) On the contrary if we reduce the load of UPR in the adaptation-alarm phase (P4), we see that the cell can recuperate using its ERAD machinery and the regulation of BiP.

delay in the onset of apoptosis (marked by slower and lower reduction of Bcl2 levels, Fig. 3.6F). This effect could be attributed to the lack of CHOP mediated branch of Bcl2 regulation. On the other hand, IRE1 and CHOP KO leads to drastic reduction in apoptosis (marked by little or no change of Bcl2 levels, Fig. 3.6F). CHOP KO, implicated the importance of CHOP in the down-regulation of Bcl2. IRE1 KO implicated the critical role of IRE1-TRAF2 mediated route of apoptosis. Overall signal flow analysis highlighted the extensive amount of crosstalk within the three branches of the UPR network. Taken together, the population of EUPR models recapitulated both short- and long timescale behavior following overload of unfolded proteins in a variety of cell types. It also captured the integration between multiple pathways and generated specific and testable hypothesis about the role of network components in signal propagation.

3.3.5 Sensitivity analysis stratified locally and globally important components of the UPR architecture.

First-order sensitivity coefficients were computed, time-averaged (over approximately 8 hrs of simulated time following ER stress) and rank-ordered for the 1090 models parameters and 636 species under normal and UPR induced conditions. Five parameter sets were selected from the parent ensemble based on rightful representation of the diversity for the sensitivity calculations (further details in the materials and methods). We selected only a few parameter sets for the sensitivity analysis, because of the computational cost of computing all 1090 coefficients over several hours of simulated time. However, we did select diverse sets as shown by the scatter in the CV values (Fig. 3.2).

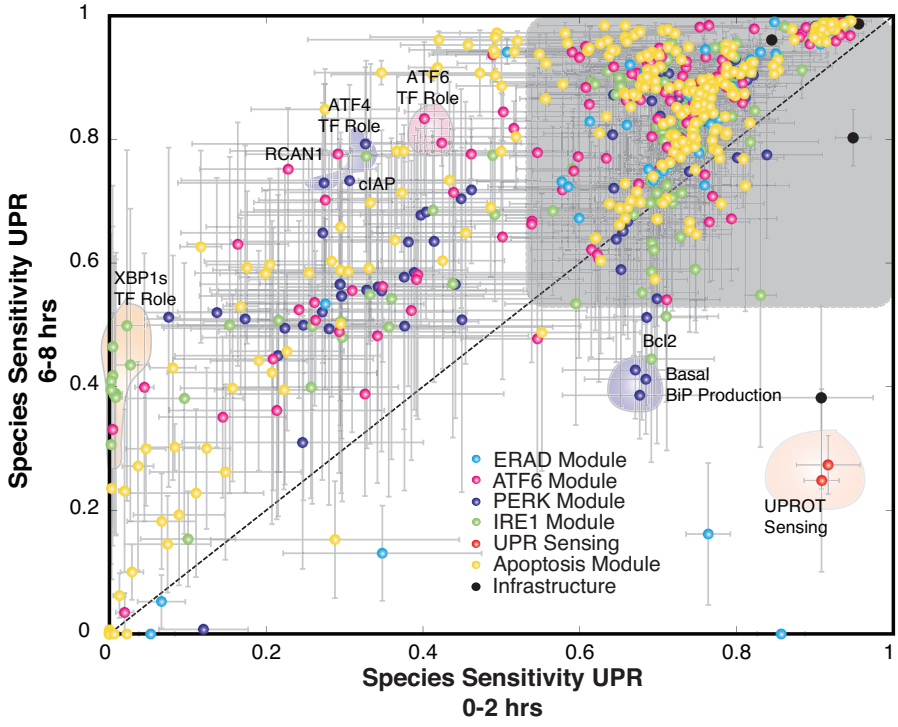


Figure 3.8: Plot of species sensitivity at earlier (0-2 hrs) versus later (6-8 hrs) time points: Sensitivity analysis was conducted over discrete two hour time windows thereby revealing the time evolution of the importance of UPR network modules. We found that signal integration via the transcriptional activity of ATF6, ATF4 and XBP1s along with RCAN1 and cIAP role in apoptosis were significantly more important at 6-8 hrs as compared to 0-2 hrs time window. This is consistent with the dominant role of the negative feedback via the transcriptional regulation of BiP in UPR. Interestingly, the majority of species rankings were similar as seen in the cluster in the grey box.

Infrastructure parameters e.g. nuclear transport, RNA polymerase or ribosome binding were globally critical, independent of stress (black points, Fig. 3.9). Additionally, apoptotic species and parameters were also important, both in the presence and absence of UPR (yellow points, Fig. 3.9). Thus, as expected, components such as RNA polymerase or caspase activation were globally important irrespective of the folding state of the ER. More interesting however were coefficients that shifted above or below the 45^o-line in the presence of

UPR. These points denote differentially important network components. While the majority of parameters and species became more important in the presence of stress, we found a band of parameters (Fig. 3.9 Inset) that were differentially important under stress. For example, the rank-ordering of the sensor and stress-transducer modules clearly increased in the presence of UPR. Approximately, 172 or 15% of the parameters were significantly more important in the presence of UPR. These parameters were largely associated with adaptation and processing of unfolded or misfolded proteins, e.g., unfolded protein degradation, cleaved ATF6-induced gene expression, IRE1-TRAF2 mediated apoptosis regulation and RCAN1 regulation. Likewise, 75 or 12% of the species were significantly more important in UPR compared with normal protein loads (data not shown).

Sensitivity analysis conducted over discrete two hour time windows revealed the time evolution of the importance of UPR network modules (Fig. 3.9). Comparison of the 0 - 2 hrs time window with itself (top panel, first column of Fig. 3.9), supported the earlier results that infrastructure components were globally critical followed by ERAD species. These species remained important in all time windows. On the other hand, during the initial 0 - 2 hrs window, ER stress transduction pathway components were robust. Comparison of the 0 - 2 hrs time window with later time points (working down the first column, Fig. 3.9), showed the increasing importance of different modules as a function of time. For example, components of the PERK and IRE1 modules were more important in the 2 - 4 hrs window compared to the earlier time points, while alarm and apoptotic phase species were more important in the 6 - 8 hrs window compared to the earlier time points. Specifically, signal integration via the transcriptional activity of ATF6, ATF4 and XBP1s along with RCAN1 and cIAP

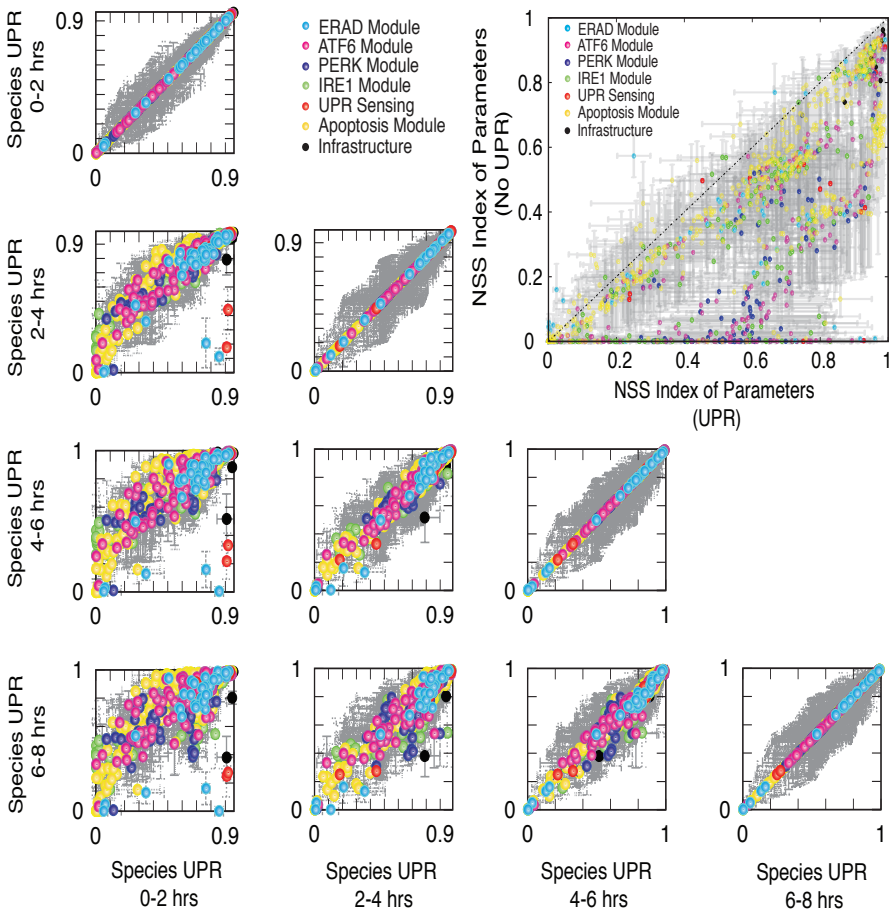


Figure 3.9: Rank-ordering of species sensitivities in the presence of UPR as a function of time. **Inset:** Rank-ordering of parameter sensitivity for UPR-induced versus normal conditions. Points denote the mean ranking computed over N = 5 parameter sets from the model population, while error bars denote one standard deviation. Points are color-coded based based upon biological function.

role in apoptosis were significantly more important at 6-8 hrs as compared to 0-2 hrs time window. This is consistent with the dominant role of the negative feedback via the transcriptional regulation of BiP in UPR. Interestingly, the majority of species rankings were similar after 6 hrs (bottom row, Fig. 3.9 and Fig. 3.8). This analysis supports the structural analysis results, wherein the signal integrators like XBP1s and transcriptional regulation of BiP via ATF6 and ATF4 are the key regulators of the phased response of UPR.

To further investigate the role of positive BiP regulation via ATF6, ATF4 and XBP1s, we conducted sensitivity analysis upon knocking out the feedback branches of BiP for the nominal parameter set (Fig. 3.10). Interestingly, upon knockout of any individual feedback branch like that of ATF4, ATF6 and XBP1s, the system overall remains equally robust. However the sensitivity of the alternate feedback components increases. This was most evident upon ATF4 feedback KO. We distinctly saw increase in sensitivity of feedback components associated with XBP1s and ATF6 (Fig. 3.10). Upon ATF6 and XBP1s feedback KO, there wasn't much change in terms of sensitivity of the system. This further attests the key regulatory effect of ATF4 in mediating the positive BiP feedback which is an essential component of the adaptation phase of UPR. Another interesting observation was that when we completely knockout all the feedback branches of BiP in the adaptation phase, the system overall becomes relatively more robust (Fig. 3.10). We distinctly saw a major shift of sensitivity of BiP upon removal of positive feedback. Overall ~ 54 % of the parameters were differentially less sensitive upon removal of BiP feedback as compared to WT. This brings to light how the presence of BiP feedback makes the system more susceptible/sensitive to perturbations. This highlights the essential role and relevance of targeting the BiP feedback in manipulating UPR and specific importance of

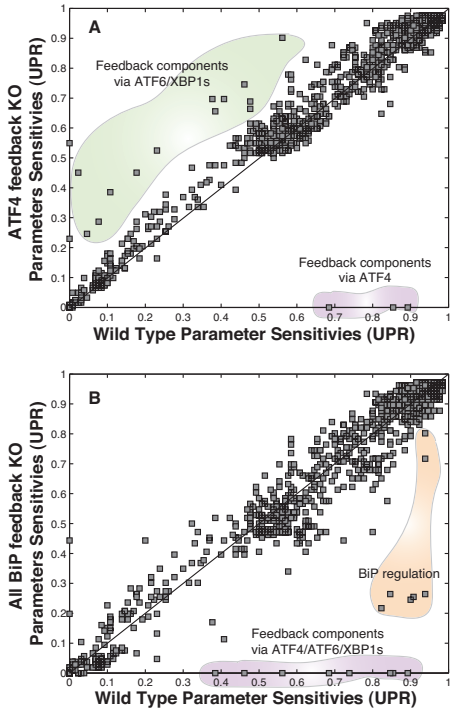


Figure 3.10: Plot of parameter sensitivity with APAF-1 feedback KO and all BiP feedback KO: Upon knockout of any individual feedback branch like that of ATF4, ATF6 and XBP1s, the system overall remains equally robust. However the sensitivity of the alternate feedback components increases. This was most evident upon ATF4 feedback KO. (A) We saw increase in sensitivity of feedback components associated with XBP1s and ATF6. Upon ATF6 and XBP1s feedback KO, there wasn't much change in terms of sensitivity of the system (data not shown). This further attests the key regulatory effect of ATF4 in mediating the positive BiP feedback which is an essential component of the adaptation phase of UPR. (B) When we completely knockout all the feedback branches of BiP in the adaptation phase, the system overall becomes relatively more robust. We distinctly saw a major shift of sensitivity of BiP upon removal of positive feedback. Overall ~ 54 % of the parameters were differentially less sensitive upon removal of BiP feedback as compared to WT. This brings to light how the presence of BiP feedback makes the system more susceptible/sensitive to perturbations.

ATF4 amongst the three feedback branches.

3.3.6 Robustness analysis predicted the phenotypic consequence of structural perturbations to the UPR network.

We calculated the direction of unfolded protein load induced concentration shifts for 636 markers following single parameter knockouts (Edge KO), single gene knockouts (GKO) and single gene overexpression (GOX) to the UPR network for the nominal parameter set (Fig. 3.11). Robustness coefficients were used to quantify the effect of structural perturbations on network markers. Coefficients with values > 1 (< 1) indicated a marker increased (decreased) compared to the basal state, while a value ~ 1 indicated approximately no change following a perturbation. Phenotypic behavior of the UPR models were analyzed as a result of perturbation, using downstream markers like Caspase 3 (marker for cell death), Bcl2 (marker for cell-survival) and other key signal integrators of the signals from ER stress transducers e.g., XBP1s, CHOP and ATF4.

Overall, we found that in the survival-death phenotypic plane, the pro-death phenotype (marked by robustness coefficients > 1 for Caspase 3) was seen to be relatively robust to structural (both GKO, GOX and Edge KO) perturbations. Few perturbations lead to increased Caspase 3 levels, e.g., overexpression of Procaspases 9/3 (Fig 3.11 A, Table 3.2). This robustness of the apoptotic marker caspase 3, can be attributed to the redundant sources of cell death (e.g., APAF-1 dependent and APAF-1 independent strategies). To confirm this we ran simulated APAF-1 KOs over the entire ensemble (Fig. 3.13). We found that there were two populations of cells in the ensemble: population 1 where APAF-1 was

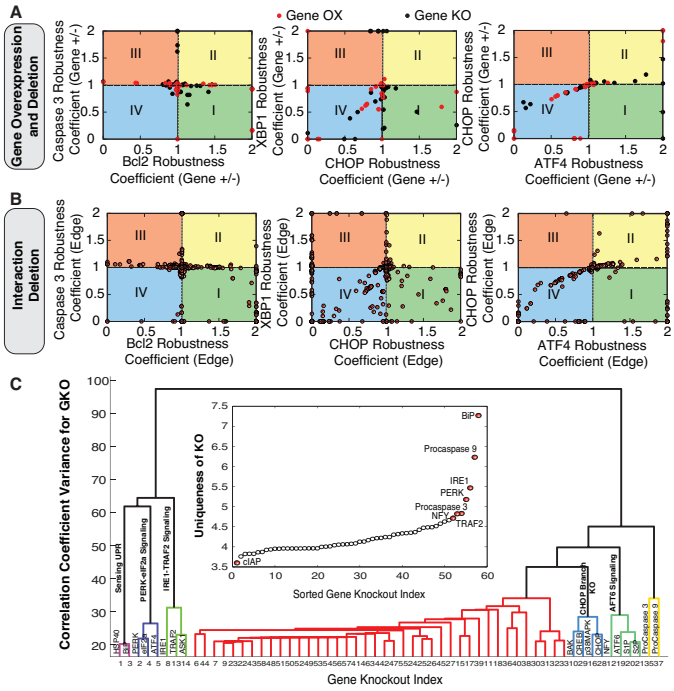


Figure 3.11: Robustness analysis of the UPR network: A-B Phenotypic phase plane analysis for the the UPR model following structural perturbations. Coupling coefficients (area under the curve from the simulation with species removed dived by the wild-type simulation) for all 636 model species were calculated for the nominal parameter set following gene overexpression/knockout (A) and deletion of single network edges (B). Coupling coefficients of one indicate no change in a marker level following a perturbation, while values less (greater) than one denote decreased (increases) marker levels. C Structural distinguishability analysis: We computed the dendrogram of the coupling coefficients for single GKO of model species. Individual coupling coefficients were clustered, where the euclidean norm was used as the distance metric and the linkage function was the inner square product (variance minimization algorithm). Each additional cluster was chosen to reduce the overall variance (y-axis). A general description of the biological function of the clusters were indicated by each group. **Insets:** Distinguishability as the magnitude of the orthogonal components for all knockout species. Species were ordered from largest to smallest magnitudes. Red markers indicate species which were statistically significant.

the dominant regulator of cell-death (marked by enhanced reduction in caspase 3 upon APAF-1 KO) and population 2 where APAF-1 is not the most dominant regulator (marked by reduced effect on Caspase 3 upon APAF-1 KO) (Fig. 3.13). This behavior is consistent based on the training data from Sak2 cells (APAF-1 $-ve$ cells) [440]. Interestingly, manipulation of the pro-survival axis via regulation of Bcl2 was possible (Fig 3.11A, Table 3.2). For example, KO of PERK/ATF4 signaling components lead to increased Bcl2 marker levels (Fig 3.11A). This behavior is attributed to the dominant role of PERK-ATF4 mediated regulation of CHOP which downstream leads to down-regulation of Bcl2 levels (seen earlier in signal analysis results). To further investigate the extent of CHOP mediated down-regulation of Bcl2, we simulated CHOP KOs over the entire ensemble (Fig. 3.13). We identified two distinct populations within the ensembles. One with a strong effect of CHOP mediated down-regulation of Bcl2 (marked by ~ 10 fold increase in Bcl2 levels) and the other with very little effect of CHOP on Bcl2 levels. This behavior could be attributed to other conflicting means of regulation of Bcl2 levels. Other schemes of regulating Bcl2 levels include perturbations involving eIF2α, BiP, JNK, cIAP and RCAN1 (Fig 3.11A, Table 3.2).

Robustness analysis allowed us to investigate complex network properties like redundancy and crosstalk. For example, the direct correlation between ATF4 and CHOP was further noted in the ATF4-CHOP phenotypic plane. Any perturbations affecting ATF4 affected CHOP levels in the same manner (reduction in ATF4 levels lead to reduced levels of CHOP) (Fig. 3.11A-B). However, owing to redundant sources of CHOP regulation (e.g., via XBP1s), effect on CHOP is damped as compared to significant changes in ATF4 levels. In the XBP1s-CHOP plane, we see at lower levels of XBP1s and CHOP, there is a direct relation between XBP1s levels and CHOP levels. However, there exists very

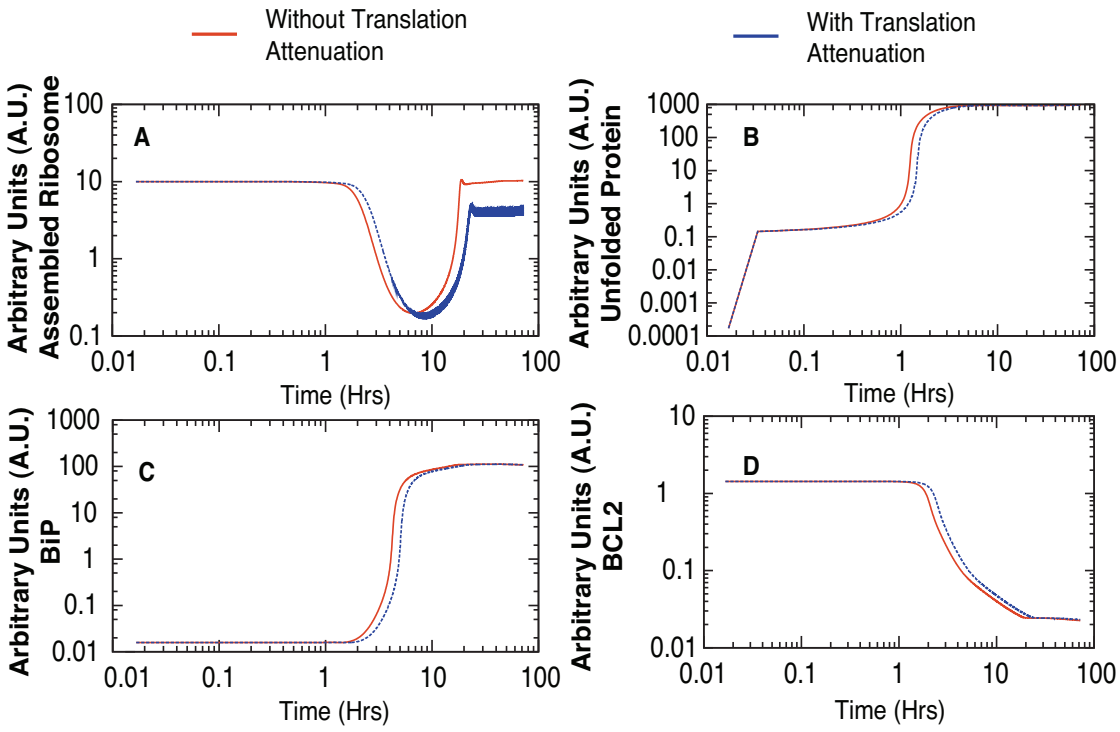


Figure 3.12: Simulations with translation attenuation built in the model: One of the key aspects which was not included in the current model was translation attenuation. So we simulated that to identify that there isnt much of a change overall in the system except for the tad bit delay in the onset of the responses.

few strategies of having both high XBP1s levels and CHOP levels indicating that higher XBP1s doesn't necessarily mean higher CHOP levels. To further investigate the implications of the feedback regulation of BiP via ATF4/ATF6/XBP1s, we simulated KOs of these components over the entire ensemble (Fig. 3.14). KO of ATF6 and XBP1s mediated feedback of BiP was seen to have little effect (as marked by robustness coefficients for BiP, Fig. 3.14). However, ATF4 mediated feedback KO led to significant amount of reduction in BiP levels (Fig. 3.14) thereby highlighting the significance of ATF4 in BiP feedback. Upon KO of all branches of BiP feedback, we found overall reductions of BiP levels. However, there were two distinct populations. One with a ~ 10 fold reduction in BiP

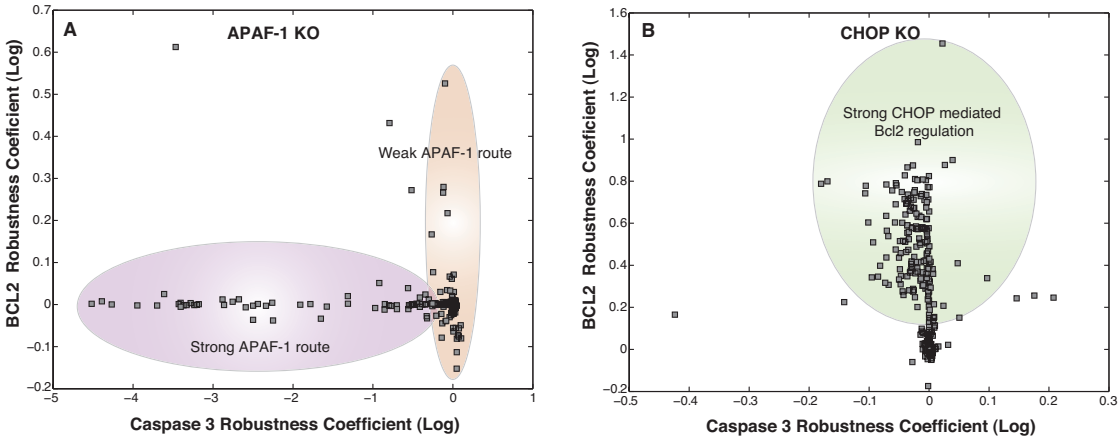


Figure 3.13: Survival-death phenotypic plane for APAF-1 and CHOP KOs over the entire ensemble: (A) With APAF-1 KO, we found that there were two populations of cells in the ensemble: population 1 where APAF-1 was the dominant regulator of cell-death (marked by enhanced reduction in caspase 3 upon APAF-1 KO) and population 2 where APAF-1 is not the most dominant regulator (marked by reduced effect on Caspase 3 upon APAF-1 KO). (B) Upon CHOP KO, we identified two distinct populations within the ensembles. One with a strong effect of CHOP mediated down-regulation of Bcl2 (marked by ~ 10 fold increase in Bcl2 levels) and the other with very little effect of CHOP on Bcl2 levels. This behavior could be attributed to other conflicting means of regulation of Bcl2 levels.

levels while the other had ~ 1000 fold reduction in BiP levels. These two populations could resemble two distinct operational paradigms within UPR. In the first mode of operation feedback regulation of BiP is really strong so when we knockout BiP feedback we have drastic reductions in BiP levels and ultimately a stronger and faster UPR response.

Global analysis of the gene knockout robustness coefficients by clustering to minimize the variance, provides a systems-level insight into the UPR network. As in Fig. 3.11C, dendrogram of the single GKOs revealed clusters with inherent functional relationships. For example, the most distinct separation was between

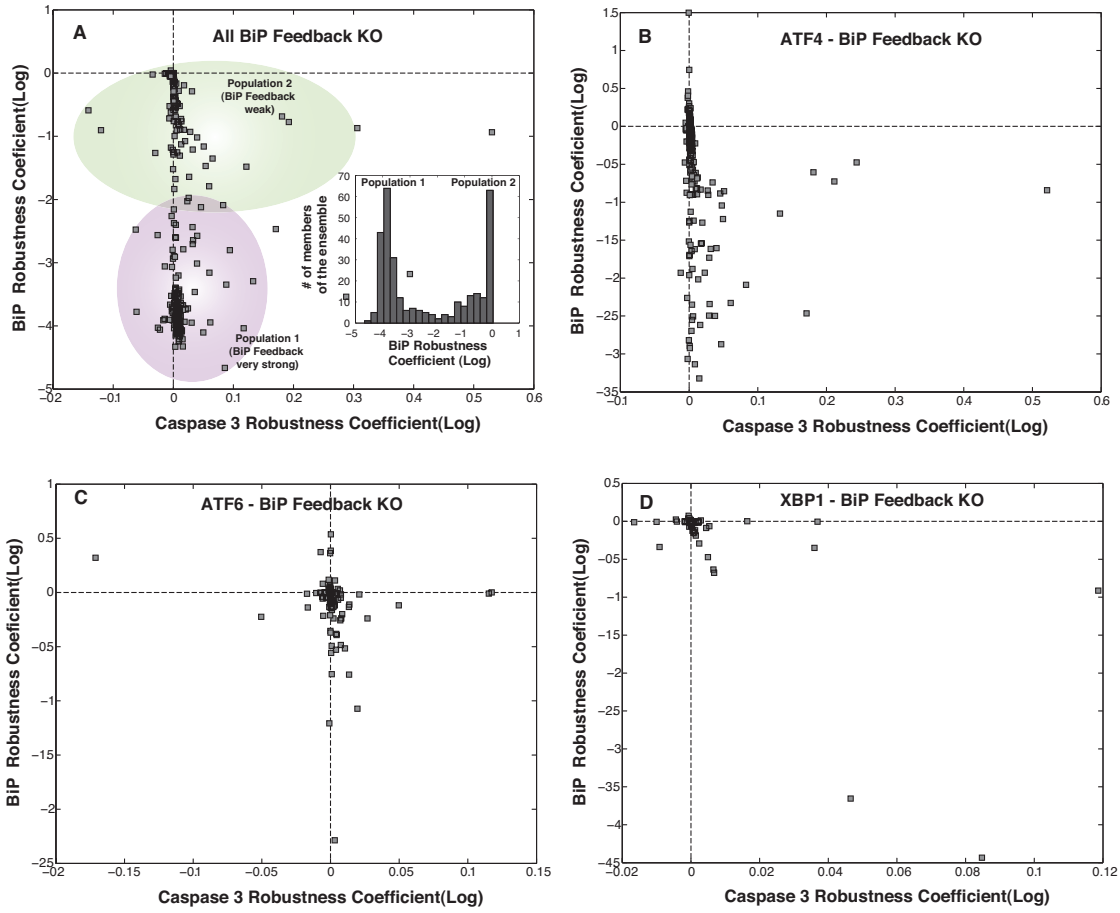


Figure 3.14: To further investigate the implications of the feedback regulation of BiP via ATF4/ATF6/XBP1s, we simulated KOs of these components over the entire ensemble. (A) Upon KO of all branches of BiP feedback, we found overall reductions of BiP levels. However, there were two distinct sub-populations. One with a $\sim$ 10 fold reduction in BiP levels while the other had $\sim$ 1000 fold reduction in BiP levels. These two populations could resemble two distinct operational paradigms within UPR. In the first mode of operation feedback regulation of BiP is really strong so when we knockout BiP feedback we have drastic reductions in BiP levels and ultimately a stronger and faster UPR response. (B) ATF4 mediated feedback KO led to significant amount of reduction in BiP levels thereby highlighting the significance of ATF4 in BiP feedback. (C-D) However, KO of ATF6 and XBP1s mediated feedback of BiP was seen to have little effect (as marked by robustness coefficients for BiP).

Table 3.2: Phenotypic response of simulated Gene knockout/overexpression. (**G.O.S** - Gene Overexpression Studies, **G.K.S** - Gene Knockout Studies)

Increased marker levels	**Phenotype**	**G.O.S (% of cases)**	**Example genes**	**G.K.S (% of cases)**	**Example genes**
BCL2	Cell Survival	16	PERK, ATF4, cIAP, Bcl2, ATF6, RCAN1, Bad, UB	24	HSP40, BiP, PERK, eIF2α, ATF4, NFY, JNK, Bad, CHOP, CREB, Procaspase 8
CHOP	–	5	eIF2α, CHOP	3.5	CEBP, CREB
Caspase 3	Cell Death	9	Procaspase 9/8/3/6/7	0	–
XBP1	–	12	BiP, cIAP, XBP1, RCAN1	3.5	NFY, TRAF2
BiP	–	12	PERK, IRE1, NFY, ATF6, CEBP, CREB, UB	12	HSP40, BiP, eIF2α, ATF4, NFY, S1P, S2P
PARP	DNA damage	3.5	Procaspase 9, PARP	0	–
ATF4	–	20	HSP40, PERK, BiP, eIF2α, ATF4, IRE1, BAK, TRAF2, ASK1, JNK, p38MAPK, Calcinuerin, CHOP, CREB	5	NFY, CEBP, CREB

unfolded protein sensing and IRE1/PERK signal initiation from the rest of the knockouts. PERK/ATF4 branch as seen earlier, plays a dominant role in the regulation of BiP and CHOP upon the onset of UPR. Similarly, IRE1/TRAF2 signaling axis is valuable to the apoptosis module. Another interesting functional module was that of CHOP, involving p38MAPK which leads to down-regulation of Bcl2 levels which considerably affects the apoptosis module. We computed the magnitude of the orthogonal components of the single GKO coefficients (Fig. 3.11C Inset). The orthogonal component was used to establish the uniqueness of the knockout. All the single GKOs were found to have orthogonal components greater than one with a 95% confidence. KO of BiP, Procaspase 9, IRE1, PERK and TRAF2 were some of the knockouts which produced the most unique effects. This is in accord to the critical role of BiP in initiating and further regulating the time scale of progression and the ultimate result of the UPR response. So PERK via ATF4 plays a key role in regulation of BiP, is rightly seen

to have a major effect as compared to other GKOs. Similarly, regulation of the apoptosis branch via IRE1-TRAF2 and Procaspase 9 rightly were seen to have the most effect as compared to other GKOs.

3.4 Discussion

Proteins requiring post-translational modifications such as N-linked glycosylation or disulfide bond formation are processed in the endoplasmic reticulum (ER). A diverse array of cellular stresses can lead to dysfunction of the ER, and ultimately to an imbalance between protein-folding capacity and protein-folding load. Unfolded or misfolded proteins are tagged for degradation via ER associated degradation (ERAD) or sent back through the folding cycle. Continued accumulation of incorrectly folded proteins can also trigger the Unfolded Protein Response (UPR). In this study, we formulated a mechanistic model of the cellular response to stress (protein overload) and studied the core regulatory aspects and downstream effects of UPR induction. PRKR-like ER kinase (PERK), inositol-requiring kinase 1 (IRE1) and activating transcription factor 6 (ATF6) were modeled as the key UPR initiators. While UPR has been extensively studied [363, 456, 245, 117, 134, 481, 167, 345, 32, 208, 554, 515], a detailed mathematical model to investigate the complexities involved is lacking. The UPR network architecture is based on extensive review of the literature [363, 456, 245, 117, 134, 481, 167, 345, 32, 208, 554, 515]. Mass balance equations describing 636 species interconnected by 1090 interactions were formulated using mass-action kinetics within an ordinary differential equation (ODE) framework. Four model populations were estimated using multi-objective optimization (33 objective functions) in conjunction with a leave-eight out cross-

validation strategy using POETs [504]. These model populations were then analyzed using population-based sensitivity and robustness analysis. Overall we identified sources of network crosstalk and redundancy within the UPR module which could be the reason for emergent properties and be the links for aberrations in cellular adaptation to stress.

A key finding of our study was that the overall outcome of UPR was as a result of simultaneous firing and competition between signaling mediated by the three ER-stress transducers: PERK, IRE1 and ATF6. This is in contrast to the traditional belief that PERK and ATF6 branches are activated before IRE1 [515]. So what we hypothesize is that instead of a sequential ordering of these branches, the state of the cell in terms of adaptation, alarm or apoptosis is a result of counteracting effects of these three prongs of UPR signaling. The counteracting/competing effects was further substantiated in simulated knockout studies, where in knockout of one ER stress transducer lead to enhancement of the other branches of UPR. Signal transduction architectures frequently contain redundancy, feedback and crosstalk. These topological features ensure signal propagation is adaptable, efficient and robust. However, they also make reprogramming signal flow challenging. This was highlighted remarkably in case of UPR. Signals from the three ER-stress transducers converged at the level of up-regulation of BiP. This is the key junction which regulates the three stages (onset and time) of adaptation/alarm and apoptosis. In this regard, regulators of this feedback cleaved ATF6, ATF4 and XBP1s were seen as highly sensitive components of UPR. Interestingly these components put-together increased the overall fragility of the system and presenting more scope of manipulation of the UPR response. This was substantiated by sensitivity analysis upon KO of the feedback loops, where we saw increased stability of the UPR module. When

these feedback components were knocked out individually, the system overall remained stable thanks to increased activity and load sharing via the other feedback branches. Amongst the three components of feedback, we identified ATF4 as the key load bearer/regulator. This was substantiated by signal flow, robustness and sensitivity analysis. This is really interesting as ATF4 protein has shown to be present greater levels in cancer compared to normal tissue, and it is up-regulated by signals of the tumor microenvironment such as hypoxia/anoxia, oxidative stress, and ER stress [10]. So any aberrations in regulation of ATF4 could potentially serve as a specific target in cancer therapy. As a target ATF4 is attractive because it is also potentially involved in angiogenesis and adaptation of cancer cells to hypoxia/anoxia, which are major problems in cancer progression [10].

Downstream effects of UPR ranges from cellular adaptation/survival (low stress) to the cell committing to apoptosis mediated death (high stress). Our modeling analysis suggested that the cell-death phenotype (marked by increased levels of Caspase 3 as compared to WT) was relatively robust. This robustness could be attributed to redundant routes of APAF-1 dependent and APAF-1 independent routes of apoptosis. This claim is supported by experimental evidence in Sak2 cells [440] and as seen by our simulated knockout studies where we identified two distinct populations representing clear distinctions in APAF-1 dependent and independent routes of apoptosis. Interestingly, manipulation of the pro-survival phenotype (marked by increased levels of Bcl2 as compared to WT) was feasible. The most effective route was via manipulation of PERK/ATF4/CHOP branch. This was substantiated by simulated CHOP KO experiments over the entire ensemble, wherein we identified two distinct populations within the ensembles. One with a strong effect of CHOP mediated

down-regulation of Bcl2 (marked by ~ 10 fold increase in Bcl2 levels) and the other with very little effect of CHOP on Bcl2 levels. This complex network behavior could be attributed to other conflicting means of regulation of Bcl2 levels. Rightfully so, induction of CHOP is involved in the development of various diseases and several therapeutic interventions [395]. For instance, suppression of CHOP by RNA interference, decoy oligodeoxynucleotides or drug inhibitors have a significant therapeutic potential to modulate type I diabetes and brain ischemia. On the other hand, overexpression of CHOP may represent a new class of anticancer therapy. Since induction of BiP has been observed in a variety of tumor cells, overexpression of CHOP directed by the BiP promoter may be used as a highly specific therapy for cancer [395]. Model analysis also highlighted the essential role of RCAN1 and IRE1-TRAF2 routes of apoptosis. ATF6 induces regulator of calcineurin 1 (RCAN1) expression [32]. RCAN1 sequesters calcineurin [32], a calcium activated protein-phosphotase B, that dephosphorylates Bcl2-antagonist of cell death (BAD) at S75 or S99 [571]. This leads to sequestering of Bcl2 by Bad, which inhibits its downstream anti-apoptotic activity [571]. Recently, a number of ATF6 homologs have been identified, e.g., OASIS, CREBH, LUMAN/CREB3, CREB4 and BBF2H7 that are processed in a similar way as ATF6, yet their function remains unknown [457]. Thus, ER-stress induced ATF6 signaling may be responsible for additional undiscovered functionality.

Given the structural and parametric uncertainty associated with the current version of the UPR model, we were still able to extract interesting insight into the complex inner workings and present falsifiable hypothesis regarding manipulating the UPR program. While we did an extensive search of the literature to formulate the model, we were missing certain key structural aspects

of UPR which could provide a more comprehensive analysis for further generations of the model. One key missing aspect is the negative regulation of the three ER-stress transducers. Given PERK's central role in translation attenuation, cells have evolved multiple axes to regulate PERK activity. First, the cytosolic kinase domain of PERK can be inhibited by the action of the DNAJ family member P58^{IPK}. P58^{IPK} was initially discovered as an inhibitor of the eIF2α protein kinase PKR [292]. P58^{IPK}, whose expression is induced following ATF6 activation, binds to the cytosolic kinase domain of PERK, inhibiting its activity [599, 558]. Inhibition of PERK kinase activity relieves eIF2α phosphorylation, thereby removing the translational block. Interestingly, P58^{IPK} expression occurs several hours after PERK activation and eIF2α phosphorylation. Thus, P58^{IPK} induction may mark the end of UPR adaptation, and the beginning of the alarm/apoptosis phase of the response [515]. Second, PERK induces a negative feedback loop, through its downstream effector CHOP, involving the direct de-phosphorylation of eIF2α. CHOP induces the expression of GADD34 which, in conjunction with protein phosphatase 1 (PP1), assembles into a phosphatase which dephosphorylates the S51 residue of eIF2α [384]. GADD34 is a member of the GADD family of genes which are induced by DNA damage and a variety of other cellular stresses [621]. The GADD34 binding partner in this complex appears to be responsible for PP1α recognition and targeting of the phosphatase complex to the ER. Association between GADD34 and PP1 is encoded by a C-terminal canonical PP1 binding motif, KVRF, while approximately 180 residues, near the N-terminus of GADD34, appear to be responsible for ER localization [56]. Currently, little is known about deactivation of ATF6. Recently, XBP1u, the unspliced form of XBP1, has been implicated as a negative regulator for ATF6 [615]. Following, the induction of ER stress, two versions of

XBP1 exist: XBP1u and sXBP1 [615]. In the recovery phase following ER stress, high levels of XBP1u may play a dual role. First, XBP1u binds sXBP1, promoting complex degradation [614, 541]. Second, XBP1u can bind ATF6α rendering it more prone to proteasomal degradation [615]. Taken together, these two steps may slow the transcription of ER chaperones and ERAD components during the recovery phase following ER stress. IRE1α activity is regulated by several proteins, including tyrosine phosphatase 1B (PTP-1B), ASK1-interactive protein 1 (AIP1) and members of the Bcl2 protein family. PTP-1B has been implicated in a number of IRE1α signaling events. The absence of PTP-1B reduced IRE1α dependent JNK activation, XBP1 splicing and EDEM transcription in immortalized and primary mouse embryonic fibroblasts [171]. However, no physical interaction between IRE1α and PTP-1B was established. On the other hand, AIP1 physically interacts with both TRAF2 and IRE1α, suggesting a model in which AIP1 facilitates IRE1α dimerization and activation [321]. The C-terminal period-like domain (PER) of AIF1 binds the N-terminal RING finger domain of TRAF2, followed by ASK1-JNK signaling [623]. Thus, based on these findings, Luo *et al.* postulated that AIF1 may be directly involved in the IRE1α-TRAF2 complex and its activation of the ASK1-JNK signaling axis [321]. This hypothesis was validated in AIP1-KO mouse studies; AIP1-knockout mouse embryonic fibroblasts and vascular endothelial cells showed significant reductions in ER-stress induced ASK1-JNK activation that was rescued in AIP1 knock-in cells [321]. IRE1α has also been shown to directly interact with Bcl-2 family members Bax and Bak. Hetz *et al.* showed that Bax and Bak complex with the cytosolic domain of IRE1α and modulate IRE1α signaling [207]. Bax and Bak double knockout mice failed to signal through the IRE1α UPR branch following tunicamycin-induced ER stress; however, PERK signaling markers, e.g., eIF2α

phosphorylation, responded normally [207]. This pro-activation role of Bak and Bax may be modulated by one of the few negative regulators of IRE1α activity, Bax inhibitor 1 (BI-1). BI-1 is an anti-apoptotic protein that enhances cell survival following several intrinsic death stimuli [593]. Bailly-Maitre *et al.* were the first to suggest that BI-1 may downregulate IRE1α and possibly ATF6 activity [21]. BI-1 deficient mice displayed increased XBP1s and enhanced JNK activity in the liver and kidney, while eIF2α phosphorylation remained normal under ER-stress conditions [21]. Lisbona *et al.* later showed that BI-1 directly interacts with the cytosolic domain of IRE1α, inhibiting its endoribonuclease activity [307]. Interestingly, BI-1 interacts with several members of the Bcl2 protein family e.g., Bcl2 and Bcl-X_L, even though it has no homology [593]. Members of the HSP family of proteins have also been shown to regulate IRE1α. For example, HSP90 interacts with the cytosolic domain of IRE1α, potentially protecting it from degradation by the proteasome [332]. HSP72 interaction with the cytosolic IRE1α domain has also recently been shown to enhance IRE1α endoribonuclease activity [175]. Taken together, these modes of IRE1α regulation with the exception of B1-1, largely promote or enhance IRE1α signaling. Given the importance of CHOP in regulation of Bcl2, it is vital to establish the exact connectivity. However, while CHOP expression is negatively correlated with Bcl2 levels, there is no CHOP binding site in the *bcl2* promoter [345]. McCullough *et al.* have suggested that the bZIP domain of CHOP could act with other bZIP transcription factors to regulate *bcl2* expression [345]. Thus, it's likely that the connection between CHOP expression and apoptosis is more complex than simple down-regulation of Bcl2 expression. These missing structural connections shall allow us to establish a detailed model and extract more relevant insights into manipulating UPR.

3.5 Materials and Methods

3.5.1 Formulation and solution of the model equations.

The unfolded protein response model was formulated as a set of coupled ordinary differential equations (ODEs):

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{r}(\mathbf{x}, \mathbf{p}) \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{3.1}$$

The symbol $\mathbf{S}$ denotes the stoichiometric matrix (636×1090). The quantity $\mathbf{x}$ denotes the concentration vector of proteins or protein complexes (636×1). The term $\mathbf{r}(\mathbf{x}, \mathbf{p})$ denotes the vector of reaction rates (1090×1). Each row in $\mathbf{S}$ described a protein or protein-protein complex, while each column described the stoichiometry of network interactions. Thus, the (i, j) element of $\mathbf{S}$, denoted by σ_{ij}, described how protein i was involved in rate j. If $\sigma_{ij} < 0$, then protein i was consumed in r_j. Conversely, if $\sigma_{ij} > 0$, protein i was produced by r_j. Lastly, if $\sigma_{ij} = 0$, there was no protein i in rate j. All of these interactions were obtained from the literature.

We assumed mass-action kinetics for each interaction in the network. The rate expression for interaction q was given by:

$$r_q\left(\mathbf{x}, k_q\right) = k_q \prod_{j \in \left\{\mathbf{R}_q\right\}} x_j^{-\sigma_{jq}} \tag{3.2}$$

The set $\left\{\mathbf{R}_q\right\}$ denotes reactants for reaction q while σ_{jq} denotes the stoichiometric coefficient (element of the matrix $\mathbf{S}$) governing species j in reaction q. All reversible interactions were split into two irreversible steps. The mass-action formulation, while expanding the dimension of the UPR model, regularized the mathematical structure; this allowed automatic generation of the

model code using UNIVERSAL and regularized the unknown model parameters (parameters were one of only three types, association, dissociation or catalytic rate constants). UNIVERSAL, an open source Java code generator, generates multiple code types from text and SBML inputs. UNIVERSAL is freely available as a Google Code project (http://code.google.com/p/universal-code-generator/). Thus, although mass-action kinetics increased the number of parameters and species, they reduced the complexity of model analysis. In this study, we considered well-mixed nuclear, cytosolic and extracellular compartments. The model equations were solved using the LSODE routine in OCTAVE (v 3.1.0; www.octave.org) on an Apple workstation (Apple, Cupertino, CA; OS X v10.6.4).

Unfolded protein response conditions were simulated by running the model to steady state and then providing a dose of proteins. The steady-state was estimated numerically by repeatedly solving the model equations and estimating the difference between subsequent time points:

$$\|\mathbf{x}(t + \Delta t) - \mathbf{x}(t)\|_2 \leq \gamma \tag{3.3}$$

The quantities $\mathbf{x}(t)$ and $\mathbf{x}(t + \Delta t)$ denote the simulated concentration vector at time t and $t + \Delta t$, respectively. The L_2 vector-norm was used as the distance metric. We used $\Delta t = 1$ s and $\gamma = 0.001$ for all simulations.

3.5.2 Estimation and cross-validation of a population of models using Pareto Optimal Ensemble Techniques (POETs).

POETs is a multiobjective optimization strategy which integrates several local search strategies e.g., Simulated Annealing (SA) or Pattern Search (PS) with a

Pareto-rank-based fitness assignment [504]. Denote a candidate parameter set at iteration $i+1$ as $\mathbf{k}_{i+1}$. The squared error for $\mathbf{k}_{i+1}$ for training set j was defined as:

$$E_j(\mathbf{k}) = \sum_{i=1}^{\mathcal{T}_j} \left(\hat{\mathcal{M}}_{ij} - \hat{y}_{ij}(\mathbf{k}) \right)^2 \tag{3.4}$$

The symbol $\hat{\mathcal{M}}_{ij}$ denotes scaled experimental observations (from training set j) while the symbol $\hat{y}_{ij}$ denotes the scaled simulation output (from training set j). The quantity i denotes the sampled time-index and $\mathcal{T}_j$ denotes the number of time points for experiment j. The read-out from the training immunoblots was band intensity where we assumed intensity was only loosely proportional to concentration. Suppose we have the intensity for species x at time $i = \{t_1, t_2, .., t_n\}$ in condition j. The scaled measurement would then be given by:

$$\hat{\mathcal{M}}_{ij} = \frac{\mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}}{\max_i \mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}} \tag{3.5}$$

Under this scaling, the lowest intensity band equaled zero while the highest intensity band equaled one. A similar scaling was defined for the simulation output.

We computed the Pareto rank of $\mathbf{k}_{i+1}$ by comparing the simulation error at iteration $i+1$ against the simulation archive $\mathbf{K}_i$. We used the Fonseca and Fleming ranking scheme [133]:

$$rank\left(\mathbf{k}_{i+1} \mid \mathbf{K}_i\right) = p \tag{3.6}$$

where p denotes the number of parameter sets that dominate parameter set $\mathbf{k}_{i+1}$. Parameter sets on or near the optimal trade-off surface have small rank. Sets with increasing rank are progressively further away from the optimal trade-off surface. The parameter set $\mathbf{k}_{i+1}$ was accepted or rejected by the SA with probability $\mathcal{P}\left(\mathbf{k}_{i+1}\right)$:

$$\mathcal{P}(\mathbf{k}_{i+1}) \equiv \exp\left\{-rank\left(\mathbf{k}_{i+1} \mid \mathbf{K}_i\right)/T\right\} \tag{3.7}$$

where T is the computational annealing temperature. The initial temperature $T_o = n/log(2)$, where n is user defined ($n = 4$ for this study). The final temperature was $T_f = 0.1$. The annealing temperature was discretized into 10 quanta between T_o and T_f and adjusted according to the schedule $T_k = \beta^k T_0$ where β was defined as:

$$\beta = \left(\frac{T_f}{T_o}\right)^{1/10} \tag{3.8}$$

The epoch-counter k was incremented after the addition of 100 members to the ensemble. Thus, as the ensemble grew, the likelihood of accepting parameter sets with a large Pareto rank decreased. To generate parameter diversity, we randomly perturbed each parameter by $\leq \pm 25\%$. We performed a local pattern-search every q steps to minimize the residual for a single randomly selected objective. The local pattern-search algorithm has been described previously [149, 560]. The parameter ensemble used in the simulation and sensitivity studies was generated from the low-rank parameter sets in $\mathbf{K}_i$.

We simultaneously calculated training and prediction error during the parameter estimation procedure using leave-eight-out cross-validation [261]. The complete set of training data (33 objectives) was subdivided into four bins; in each bin 25 data sets were reserved for training while eight were reserved for prediction. In the first bin $DS_1 \ldots DS_8$ were used for validation while $DS_9 \ldots DS_{33}$ were used for training. In the second bin $DS_9 \ldots DS_{16}$ were used for validation while $DS_1 \ldots DS_8$ $DS_{17} \ldots DS_{33}$ were used for training, etc. Thus, we formulated four ensembles from which we evenly selected parameter sets for the *parent* ensemble (Fig. 3.2). While cross-validation required that we generate additional model populations, we trained and tested against all the data sets.

3.5.3 Sensitivity and robustness analysis of the population of UPR models.

Sensitivity coefficients were calculated as shown previously [504] using five models selected from the ensemble (red points, Fig. 3.2). The resulting sensitivity coefficients were scaled and time-averaged (Trapezoid rule):

$$\mathcal{N}_{ij} \equiv \frac{1}{T} \int_0^T dt \cdot |\alpha_{ij}(t)\, s_{ij}(t)| \tag{3.9}$$

where T denotes the final simulation time and $\alpha_{ij} = 1$. The time-averaged sensitivity coefficients were then organized into an array for each ensemble member:

$$\mathcal{N}^{(\epsilon)} = \begin{pmatrix} \mathcal{N}_{11}^{(\epsilon)} & \mathcal{N}_{12}^{(\epsilon)} & \dots & \mathcal{N}_{1j}^{(\epsilon)} & \dots & \mathcal{N}_{1P}^{(\epsilon)} \\ \mathcal{N}_{21}^{(\epsilon)} & \mathcal{N}_{22}^{(\epsilon)} & \dots & \mathcal{N}_{2j}^{(\epsilon)} & \dots & \mathcal{N}_{2P}^{(\epsilon)} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathcal{N}_{M1}^{(\epsilon)} & \mathcal{N}_{M2}^{(\epsilon)} & \dots & \mathcal{N}_{Mj}^{(\epsilon)} & \dots & \mathcal{N}_{MP}^{(\epsilon)} \end{pmatrix} \qquad \epsilon = 1, 2, \dots, N_\epsilon \tag{3.10}$$

where ϵ denotes the index of the ensemble member, P denotes the number of parameters, N_ϵ denotes the number of ensemble samples and M denotes the number of model species. To estimate the relative fragility or robustness of species and reactions in the network, we decomposed the $\mathcal{N}^{(\epsilon)}$ matrix using Singular Value Decomposition (SVD):

$$\mathcal{N}^{(\epsilon)} = \mathbf{U}^{(\epsilon)} \Sigma^{(\epsilon)} \mathbf{V}^{T,(\epsilon)} \tag{3.11}$$

Coefficients of the left (right) singular vectors corresponding to largest β singular values of $\mathcal{N}^{(\epsilon)}$ were rank-ordered to estimate important species (reaction) combinations. Only coefficients with magnitude greater than a threshold (δ = 0.1) were considered. The fraction of the β vectors in which a reaction or species index occurred was used to rank its importance.

Robustness coefficients of the form:

$$\alpha\left(i, j, t_o, t_f\right) = \left(\int_{t_o}^{t_f} x_i\left(t\right) dt\right)^{-1} \left(\int_{t_o}^{t_f} x_i^{(j)}\left(t\right) dt\right) \tag{3.12}$$

were calculated to understand the robustness of the network. The robustness coefficient $\alpha\left(i, j, t_o, t_f\right)$ is the ratio of the integrated concentration of a network marker in the presence (numerator) and absence (denominator) of structural or operational perturbation. The quantities t_0 and t_f denote the initial and final simulation time respectively, while i and j denote the indices for the marker and the perturbation respectively. If $\alpha\left(i, j, t_o, t_f\right) > 1$, then the perturbation *increased* the marker concentration. Conversely, if $\alpha\left(i, j, t_o, t_f\right) \ll 1$ the perturbation *decreased* the marker concentration. Lastly, if $\alpha\left(i, j, t_o, t_f\right) \sim 1$ the perturbation did not influence the marker concentration.

3.5.4 Species clustering and dendrogram.

A dendrogram was derived by considering each of the knockouts(over-expressions) as variables and the average log of robustness coefficient (LRC) for each of the species as observations. We used the Euclidean norm in LRC space as the distance metric. The linkage function (objective function for identifying variable clusters) was the inner squared distance (minimum variance algorithm). The Statistical Toolbox of Matlab (The Mathworks, Natick, MA) was used to generate the distances, linkages and the final dendrogram.

3.5.5 Identification of distinguishable species.

Robustness coefficients were used to rank-order knockout(overexpression) experiments in terms of the greatest unique responses and identify species which were linearly distinguishable. The response of the knockout(overexpression) was measured in terms of the robustness coefficients. The LRC had desirable linear properties, such that no response (no change in trajectories from wild-type) returns a value of zero and similar negative and positive responses have different directions but similar magnitudes. We considered the unique component of the response to be the orthogonal component in LRC space and the magnitude of the response to be the Euclidean norm. The orthogonal components and there magnitude were identified for each parameter set in the ensemble by first choosing the knockout(overexpression) with the greatest magnitude, x_1 and placing it in the empty set $\mathcal{V}$. The knockout(overexpression) x_1 defines the orthogonal directions in the LRC space. We then calculated the orthogonal components for all remaining knockouts(overexpressions) relative to x_1, and added the knockout(overexpression) species with the greatest orthogonal magnitude to set $\mathcal{V}$. In general the components of all remaining x_i orthogonal to set $\mathcal{V}$ were calculated and the largest was moved into set $\mathcal{V}$. This process was continued until all knockout(overexpression) species, x_i were added to set $\mathcal{V}$. Mathematically two species were considered distinguishable if and only if they were linearly independent (the orthogonal components were non-zero). We considered a threshold value of one or five and performed a student t-test (Matlab Statistical Toolbox, The Mathworks, Natick, MA) to identify which species had orthogonal components above the threshold with a 95% confidence over the ensemble.

CHAPTER 4

COMPUTATIONAL MODELING AND ANALYSIS OF INSULIN INDUCED EUKARYOTIC TRANSLATION INITIATION

Authors - Lequieu JP, Chakrabarti A, Nayak S and Varner JD.

Published in - PLoS Comput Biol 7(11): e1002263.

4.1 Abstract

Insulin, the primary hormone regulating the level of glucose in the bloodstream, modulates a variety of cellular and enzymatic processes in normal and diseased cells. Insulin signals are processed by a complex network of biochemical interactions which ultimately induce gene expression programs or other processes such as translation initiation. Surprisingly, despite the wealth of literature on insulin signaling, the relative importance of the components linking insulin with translation initiation remains unclear. We addressed this question by developing and interrogating a family of mathematical models of insulin induced translation initiation. The insulin network was modeled using mass-action kinetics within an ordinary differential equation (ODE) framework. A family of model parameters was estimated, starting from an initial best fit parameter set, using 24 experimental data sets taken from literature. The residual between model simulations and each of the experimental constraints was simultaneously minimized using multiobjective optimization. Interrogation of the model population, using sensitivity and robustness analysis, identified an insulin-dependent switch that controlled translation initiation. Our analysis suggested that without insulin, a balance between the pro-initiation activity of the GTP-binding

protein Rheb and anti-initiation activity of PTEN controlled basal initiation. On the other hand, in the presence of insulin a combination of PI3K and Rheb activity controlled inducible initiation, where PI3K was only critical in the presence of insulin. Other well known regulatory mechanisms governing insulin action, for example IRS-1 negative feedback, modulated the relative importance of PI3K and Rheb, but did not fundamentally change the signal flow.

4.2 Introduction

Insulin, the primary hormone regulating the level of glucose in the bloodstream, modulates a variety of cellular and enzymatic processes in normal and diseased cells [636, 358, 89, 462, 477, 619, 118]. The regulation of cellular function by insulin and insulin-like growth factors I/II (IGF-I/II) is a highly complex process [413, 128, 525, 467, 466, 450, 322]. Insulin and IGF-I/II interact with insulin receptors (IR), and type I/II IGF receptors (IGF-IR/IIR) in addition to other transmembrane receptors [525]. These interactions ultimately induce gene expression programs or other processes such as translation initiation. Translation rates of many cell cycle and survival proteins are modulated by growth factor, hormone or other mitogenic signals [34]. Insulin induces the activation of class I Phosphoinositide 3-kinases (PI3Ks), which in turn activate the serine/threonine protein kinase Akt and the mammalian target of rapamycin (mTOR). The PI3K/Akt/mTOR signaling axis is important to a variety of cellular programs, including apoptosis [602], cell size control [150] and translation initiation. Among other functions, activation of the PI3K/Akt/mTOR axis results in the phosphorylation of eukaryotic translation initiation factor 4E-binding protein (4E-BPx) family members [161]. Phosphorylation of 4E-BPx

causes the release of the eukaryotic translation initiation factor 4E (eIF4E), which is critical to directing ribosomes to the 7-methyl-guanosine cap of eukaryotic mRNAs. Previously, the availability of eIF4E has been shown to be rate limiting for translation initiation in many eukaryotic cell-lines [635, 34]. Given its central role in cell biology, evolutionarily optimized infrastructure like translation might be expected to be robust or highly redundant. Surprisingly, deregulated translation, especially involving growth-factor or insulin induced initiation mechanisms, has been implicated in a spectrum of cancers [42].

Despite the wealth of literature on insulin signaling, the relative importance of the components linking insulin with translation initiation remains unclear. Many investigators have explored this question using both experimental and computational tools. For example, Caron *et al.* recently published a comprehensive map of the mTOR signaling network, including a detailed portrait of insulin induced mTOR activation and its downstream role in translation initiation [68]. Taniguchi *et al.* proposed three criteria to identify the critical nodes of insulin signaling: network divergence, degree of regulation and potential crosstalk [525]. Using these criteria, they identified insulin-receptor (IR), PI3K and Akt as the *critical nodes* of insulin action. Several insightful mathematical models of insulin-signaling have also been published [122, 107, 46, 223]. While these models vary in their focus and biological scope, none has exclusively focused on how insulin stimulates translation initiation. This particular question was addressed by Nayak *et al.*, who analyzed a family of detailed mathematical models of growth factor and insulin induced translation initiation [368]. Like the Taniguchi *et al.* hypothesis, their study suggested that Akt/mTOR were structurally fragile, and likely the key elements integrating growth factor signaling with translation. However, the Nayak *et al.* model neglected several

key features of insulin processing, e.g., negative feedback of IR resulting from mTOR activity.

The objective of this study was to rank-order the importance of components of insulin-induced translation initiation using computational tools. Toward this objective, we analyzed an ensemble of mechanistic mathematical models of insulin induced translation initiation that was a significant extension of our previous work [368]. First, we expanded the original model connectivity to include a detailed description of the regulation and activity of insulin, insulin-like growth factor and platelet-derived growth factor (PDGF) receptor family members (including negative feedback). Second, we refined the description of the phosphorylation state of Akt and its downstream role in the activation of the mTORC1 and mTORC2 complexes. Lastly, we used new model estimation and interrogation techniques to generate and analyze an uncorrelated population of initiation models that were simultaneously consistent with 24 qualitative and quantitative data sets. Interrogation of this model population, using sensitivity and robustness analysis, identified an insulin-dependent switch that controlled translation initiation. Without insulin, a balance between the pro-initiation activity of the GTP-binding protein Rheb and anti-initiation activity of PTEN controlled basal initiation. Rheb knockdown simulations confirmed decreased initiation in the majority of the model population, while translation initiation increased for all models in the population following a PTEN deletion. On the other hand, a combination of PI3K and Rheb activity controlled insulin inducible initiation. PI3K deletion in the presence of insulin removed the ability of the network to process insulin signals, but did not remove initiation altogether. PI3K deletion reduced initiation to approximately 60% of its maximum level. Interestingly, the relative contribution of PI3K versus Rheb to the overall initiation level could

be tuned by controlling IRS-1 feedback. In the absence of feedback, PI3K was more important than Rheb to signal propagation, while the opposite was true in the presence of feedback. Taken together, our modeling study supported the Taniguchi *et al.* hypothesis that PI3K was a critical node in the insulin-induced initiation network. However, we also found that the role of PI3K was nuanced; PI3K in combination with Rheb controlled initiation in the presence of insulin, while the combination of PTEN and Rheb controlled basal initiation.

4.3 Results

4.3.1 Translation initiation model connectivity.

The translation initiation model consisted of 250 protein, lipid or mRNA species interconnected by 573 interactions (Fig. 4.1). The model described the integration of insulin and growth-factor signaling with 80S assembly. While other eukaryotic translation initiation mechanisms exist, we focused only on cap-mediated translation as the dominant translation mechanism [333]. The model interactome was taken from literature (SBML file available in the supplemental materials); the connectivity of insulin- and growth-factor induced translation initiation has been extensively studied [322, 228]. The model interactome was not specific to a single cell line. Rather, it was a canonical representation of the pathways involved in insulin and growth-factor induced initiation. Using a canonical network allowed us to explore general features of insulin or growth-factor induced translation initiation without cell line specific artifacts. Binding of insulin or IGF-I/II with IR or IGF-I/IIR promotes the autophosphorylation

of the cytosolic domains of these receptors at tyrosine residues. Receptor autophosphorylation promotes the formation of adaptor complexes, which are anchored in place by insulin receptor substrate (IRSx) family members; IRSx are required for the assembly of adaptor complexes involving the SHC-transforming protein 1 (Shc), Son of Sevenless (SoS), growth factor receptor-bound protein 2 (Grb2) and Ras proteins [502, 501, 452]. In the model we considered only the IRS-1 protein and neglected other IRSx family members. Adaptor complex formation ultimately culminates in the activation of the catalytic subunit of PI3K. Among their many roles, PI3Ks catalyze the phosphorylation of the phospholipid PIP2 to PIP3 [619]. PIP3 is critical to the localization of 3-phosphoinositide-dependent kinase 1 (PDK1) to the membrane, where it phosphorylates the master kinase Akt at Thr308 [357]. Akt is further phosphorylated at Ser473 by the rictor-mammalian target of rapamycin (mTORC2) protein [472]. Once phosphorylated, Akt promotes translation initiation by directly or indirectly activating the mTORC1 protein [636]. Akt directly activates mTORC1 through a novel binding partner known as PRAS40 [469, 177]. However, mTORC1 can also be activated by the GTP bound form of the Ras homologue enriched in brain (Rheb) protein. Without insulin, Rheb is regulated by the tuberous sclerosis complex TSC1/2, which has GTPase activating protein (GAP) activity. Akt directly phosphorylates TSC1/2 which inhibits its GAP activity and allows Rheb-mediated activation of mTORC1 [225, 331]. Activated mTORC1 plays two key roles in translation initiation; first, it activates ribosomal protein S6 kinase beta-1 (S6K1) and second it phosphorylates eukaryotic translation initiation factor 4E-binding protein (4E-BPx) family members [189]. In this study, we included only 4E-BP1 and modeled a single deactivating phosphorylation site. Phosphorylated 4E-BP1 releases eIF4E which, along with other initiation factors, is critical

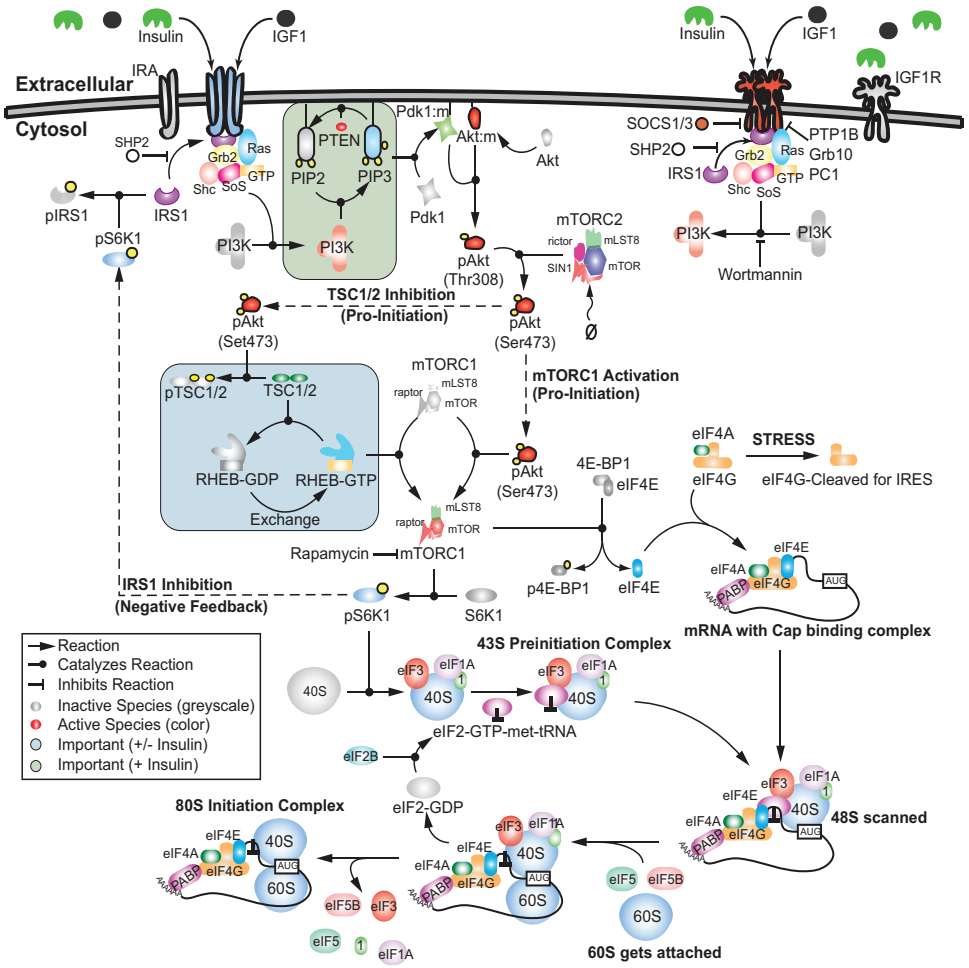


Figure 4.1: Schematic of the translation initiation signaling network. Growth factors trigger receptor dimerization and the formation of adaptor complexes which activate PI3K. PI3K then signals through PIP2/3 to activate Akt. Activated Akt can then activate mTORC1 either directly or by phosphorylating TSC1/2, an inhibitor of Rheb. Activated mTORC1 can phosphorylate 4EBP1 and activate S6K1, two necessary checkpoints for translation initiation. mTORC1 can also phosphorylate IRS-1, a negative feedback which inhibits formation of the adaptor complex and attenuates insulin signaling.

to directing ribosomes to the 7-methyl-guanosine cap structure of eukaryotic mRNAs [228].

Several mechanisms attenuate insulin and growth-factor induced transla-

tion initiation. First, insulin signal propagation can be controlled by disrupting adaptor complex formation. For example, we included tyrosine phosphatases and competitive inhibitors such as protein-tyrosine phosphatase 1B (PTP1B), src homology phosphotyrosyl phosphatase 2 (SHP2), growth factor receptor-bound protein 10 (Grb10) and suppressor of cytokine signaling 1/3 (SOCS1/3) which interfere with adaptor complex formation and activity [525, 113, 361, 549]. Second, several mechanisms control PIP3 formation, PDK1 recruitment and Akt phosphorylation [525]. In the model, we included the phosphatase and tensin homolog (PTEN) protein, which dephosphorylates PIP3 [513], as well as the SH2 (Src homology 2)-containing inositol phosphatase-1 (SHIP1) protein which hydrolyses the 5'-phosphates from PIP3 [306]. Lastly, S6K1 inhibits IRS-1 activity by phosphorylation at Ser318 [173]. S6K1/IRS-1 feedback has been shown to be important in insulin resistance and cancer [322, 551, 550, 111].

4.3.2 Estimating an ensemble of translation initiation models using POETs.

Translation initiation was modeled using mass-action kinetics within an ordinary differential equation (ODE) framework. ODEs and mass-action kinetics are common methods of modeling biological pathways [479, 74, 203]. However, ODEs have several important limitations that could be addressed with other model formulations e.g., Partial Differential Equation (PDE) based models. PDEs naturally describe spatially distributed intracellular processes or can be used to model population dynamics using population balance methods [497]. However, the computational burden associated with solving and analyzing sys-

tems of PDEs, especially at the scale of the current study, would be substantial. Alternatively, we have addressed both of these ODE shortcomings (without resorting to a PDE formulation) by including well-mixed compartments to account for spatially localized species and processes and have considered an ensemble of models in our analysis to coarse-grain population phenomena. Irregardless of whether we have an ODE or PDE model formulation, both classes of model typically require the identification of a large number of unknown model parameters. The initiation model had 823 unknown parameters (573 kinetic parameters and 250 initial conditions), which were not uniquely identifiable (data not shown). We estimated an experimentally constrained population of parameters using multiobjective optimization. Model parameters were estimated, starting from an initial best fit parameter set, using 24 *in vitro* and *in vivo* data sets taken from literature (Table 4.1). These training data were taken from multiple independent studies (in different cell lines) exploring insulin and IGF-I/II signaling or *in-vitro* translation initiation. These data were largely western blot measurements of the total or phospho-specific abundance of proteins following the addition of a stimulus or inhibitor. While the use of multiple cell-lines was not ideal, it did allow us to capture a consensus picture of insulin or IGF-I/II initiated signaling (which was useful in understanding the general operational principles of the network). However, one should be careful when applying consensus models to specific cell lines or tissues, as these generally may behave qualitatively differently.

The residual between model simulations and each of the experimental constraints was simultaneously minimized using the multiobjective POETs algorithm [504]. We used a leave-three-out cross validation strategy to independently estimate prediction and training error during parameter identification

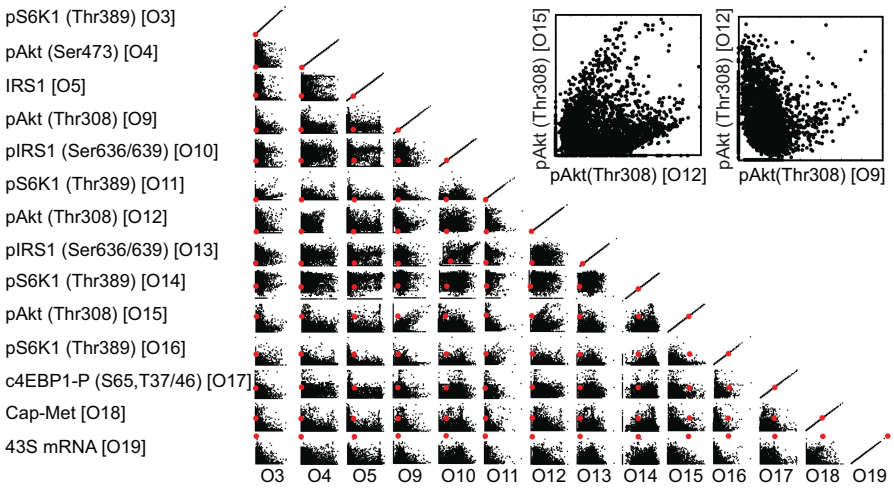


Figure 4.2: The scaled simulation error (SSE) for selected objective function pairs for N = 5818 rank zero initiation models. The SSEs for objective functions chosen by cross-validation for prediction was set to zero and disregarded when ranking other sets. The red point denotes the performance of the nominal parameter set.

(Table 4.1). Additionally, a random control (100 random parameter sets) was run to check the training/prediction fitness above random (Table 4.1). The *training* error for 23 of the 24 objectives was statistically significantly better than the random control at a 95% confidence level. Additionally, for 20 of the 24 objectives, the model *prediction* error was also significantly better than the random control ($p \leq 0.05$). Of the four remaining objectives (O4,O5,O12 and O13), three involved phosphorylated Akt (O4 and O12) or IRS-1 (O13), each of which had redundant measurements in the objective set that were significant. While the remaining objective, which involved IRS-1 levels (O5), was not significantly better than the random control, the absolute error was small.

The ensemble of translation models recapitulated diverse training data across multiple cell lines. POETs generated 18,886 probable models with Pareto rank ≤ 4. Model parameters had coefficients of variation (CV) ranging from

0.65 to 1.10. Further, 89% (512 of 573) of the model parameters were constrained with a CV < 1. The performance of 5,818 rank-zero models is shown in Fig. 4.2. The majority of objective functions were uncorrelated e.g., O4×O13 or O12×O13 or directly proportional e.g., O3×O11 or O9×O15. Uncorrelated or proportional objectives suggested the model population simultaneously described each training constraint. However, several other objectives were inversely proportional e.g., O12×O14. For these pairs, the model was unable to simultaneously fit both training data sets. Surprisingly, these objectives were the same protein pAkt(Thr308) O9×O12 and pS6K1(Thr389) O3×O14, taken from either different cell lines or different labs.

This suggested conflicts in the data e.g., cell line variation or differences in specific laboratory protocols, rather than structural inaccuracies in the model, were responsible for the inverse relationship. The key indicators of eukaryotic translation initiation are the phosphorylation of S6K1 and 4E-BP1 [189]. Both Tzatos *et al.* and Villalonga *et al.* performed insightful studies exploring the dynamics of S6K1 and 4E-BP1 phosphorylation in L6 Myotubes and RhoE 3T3 cells [548, 566]. The ensemble recapitulated these observations with error distributions that were statistically significantly better than random parameters (ϵ_{train} = 0.27±0.33, ϵ_{rand} = 2.95±0.51; ϵ_{train} = 0.57±0.33, ϵ_{rand} = 1.76±0.43) (Fig. 4.3A and 4.3B, Table 4.1). The model population also recapitulated IGF1 induced Akt and S6K1 phosphorylation (ϵ_{train} = 0.12±0.13, ϵ_{rand} = 0.69±0.11; ϵ_{train} = 0.18±0.07, ϵ_{rand} = 0.69 ± 0.11) (Fig. 4.3E and 4.3F, Table 4.1). Lorsh *et al.* studied ribosomal assembly dynamics in rabbit reticulocytes, suggesting the formation of the eIF2:GTP:Met-tRNA tertiary complex was rate limiting in 80S formation [316]. Our model captured 80S assembly dynamics, including the crucial lag phase in the first two minutes of stimulation (ϵ_{train} = 0.42 ± 0.46, ϵ_{rand} = 1.24 ± 0.71) (Fig.

Table 4.1: Objective function list along with species, cell type, cellular compartment, nominal error, training error, prediction error, random error with a randomly generated parameter set and the corresponding literature reference.

O#	Species	Cell Type	Nominal	Training	Prediction	Random	Source
O1	PI3K Activity	3T3-L1 cells	0.01	0.01 ± 0.00	0.01 ± 0.00	0.67 ± 0.20	[298]
O2	PIP3	3T3-L1 cells	0.00	0.00 ± 0.00	0.00 ± 0.00	0.84 ± 0.08	[298]
O3	pS6K1 (T389)	3T3-L1 cells	0.39	0.17 ± 0.15	0.27 ± 0.24	1.55 ± 0.49	[177]
O4	pAkt (S473)	3T3-L1 cells	0.38	0.30 ± 0.23	0.53 ± 0.29	0.50 ± 0.38	[177]
O5	IRS1	3T3-L1 cells	0.43	0.47 ± 0.62	1.37 ± 0.71	0.56 ± 0.58	[177]
O6	pAkt (S473)	393T cells	0.06	0.28 ± 0.32	0.43 ± 0.35	1.10 ± 0.31	[177]
O7	pAkt (S473)	C2C12 myotubes	0.05	0.12 ± 0.13	0.12 ± 0.13	0.69 ± 0.11	[283]
O8	pS6K1 (T421/S424)	C2C12 myotubes	0.20	0.18 ± 0.07	0.20 ± 0.10	0.47 ± 0.22	[283]
O9	pAkt (T308)	HUVEC cells	1.21	0.78 ± 0.38	0.94 ± 0.36	1.20 ± 0.79	[204]
O10	IRS-1P (S636/639)	L6 Myotubes	1.34	1.17 ± 0.37	1.13 ± 0.35	1.28 ± 0.38	[548]
O11	pS6K1(T389)	L6 Myotubes	0.98	0.27 ± 0.33	0.55 ± 0.64	2.95 ± 0.51	[548]
O12	pAkt (T308)	L6 Myotubes	0.93	0.62 ± 0.36	0.71 ± 0.34	0.84 ± 0.48	[548]
O13	IRS-1P (S636/639)	L6 Myotubes	1.24	1.07 ± 0.38	1.29 ± 0.31	1.35 ± 0.36	[548]
O14	pS6K1 (T389)	L6 Myotubes	2.36	2.02 ± 0.43	2.26 ± 0.24	1.95 ± 0.38	[548]
O15	pAkt (T308)	L6 Myotubes	0.97	0.39 ± 0.35	0.48 ± 0.33	0.87 ± 0.82	[548]
O16	pS6K1(T389)	RhoE 3T3 cells	1.33	0.28 ± 0.33	0.21 ± 0.25	2.94 ± 0.54	[566]
O17	c4EBP-P (S65, T37/46)	RhoE 3T3 cells	0.37	0.57 ± 0.33	0.85 ± 0.38	1.76 ± 0.43	[566]
O18	Cap-Met-Puro	rabbit reticulocytes	0.46	0.42 ± 0.46	0.86 ± 0.73	1.24 ± 0.71	[316]
O19	43S-mRNA	rabbit reticulocytes	0.19	0.37 ± 0.39	0.57 ± 0.47	1.14 ± 0.64	[316]
O20	pAkt (S473)	A14 NIH 3T3 cells	1.12	0.98 ± 0.23	0.99 ± 0.23	1.16 ± 0.15	[151]
O21	pS6K1(T389)	A14 NIH 3T3 cells	1.20	0.57 ± 0.29	0.57 ± 0.23	0.69 ± 0.21	[151]
O22	Rheb	HeLa cells	0.00	0.15 ± 0.83	0.10 ± 0.71	1.99 ± 0.09	[151]
O23	pS6K1(T389)	HeLa cells	0.13	0.14 ± 0.11	0.24 ± 0.23	0.77 ± 0.58	[151]
O24	c4EBP1-P (T70)	HEK293 cells	0.25	0.34 ± 0.26	0.62 ± 0.41	0.90 ± 0.22	[151]

4.3C, Table 4.1). Inhibitor data was also used for model training. Without insulin, PI3K was not activated and pAkt (Ser473) levels remained low (Fig 4.3D, lane 1). Following insulin stimulation, PI3K activation resulted in increased pAkt(Ser473) levels (Fig 4.3D, lane 2). Wortmannin, a PI3K inhibitor, significantly decreased pAkt(Ser473) (Fig 4.3D, lane 3). While our model population qualitatively captured this decrease, the levels of pAkt(Ser473) were higher than

those observed experimentally. The model was not trained using mTORC1/2 measurements, however species immediately upstream and downstream of mTORC1/2, namely pAkt(Ser473) or S6K1 were used in model training. Without insulin, pAkt(Ser473) and S6K1(Thr421/Ser424) levels were low (Fig 4.3E/F, lanes 1). Addition of insulin increased pAkt(Ser473) and S6K1(Thr421/Ser424). Upon rapamycin addition, mTORC1 was inhibited and the levels of phosphorylated S6K1 decreased (Fig. 4.3E, lane 3). However, because of its position upstream of mTORC1, pAkt(Set473) levels were unchanged (Fig. 4.3E, lane 3).

The model was validated by comparing simulations with *in vivo* and *in vitro* data sets *not* used for training or cross-validation (Table 4.2). For four of the five prediction data sets, the model demonstrated errors statistically significantly better than a random control ($p \leq 0.05$). However, the remaining prediction case (P3), while not significantly different than random, has a small error relative to the other objectives. Data from Lorsh *et al.* was used to validate the dynamics of intermediate ribosomal complexes [316]. The level of 43S mRNA was quantified using both GTP and a non-degradable GTP-like homologue GMP-PNP (Fig. 4.4A). Data involving GMP-PNP was used for training while data involving GTP was used only for validation ($\epsilon_{pred} = 0.52 \pm 0.40$, $\epsilon_{rand} = 0.82 \pm 0.51$). Garami *et al.* explored insulin-induced Rheb activation and the role of TSC1/2 in the presence and absence of wortmannin and rapamycin [151]. We first compared measured versus simulated Rheb-GTP levels, with and without insulin, in the absence of inhibitors.

While we captured the qualitative trends, we over-predicted the percentage of GTP bound Rheb ($\epsilon_{pred} = 0.22 \pm 0.11$, $\epsilon_{rand} = 0.42 \pm 0.01$) (Fig. 4.4B). The model also failed to predict sustained Rheb-GTP levels in the presence of ra-

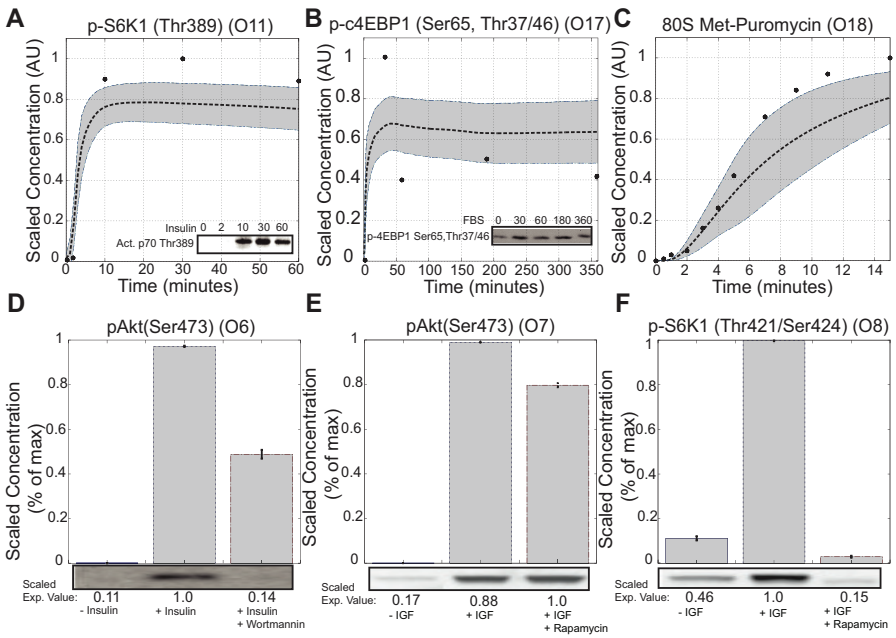


Figure 4.3: Ensemble performance against selected training objectives (N = 400). Dotted lines represent the simulation mean of the ensemble, while the shaded region denotes the 99.9% confidence estimate for the mean. The solid dots represent the scaled experimental data. **A.** Time course data for p70S6K1 phosphorylation in response to insulin stimulation (L6 Myotubes). **B.** Time course data for c4EBP1 phosphorylation in response to FBS (RhoE 3T3 cells). **C.** In vitro time course of the 80S complex measured by puromycin assay (rabbit reticulocyte). **D.** pAkt(Ser473) levels at 20 minutes in the presence and absence of insulin and wortmannin (393T cells). **E,F.** pAkt(Set473) and activated p70S6K1 levels at 15 minutes in the presence and absence of insulin-like growth factor (IGF) and rapamycin (C2C12 myotubes).

pamycin. This suggested that sustained pAkt(Ser473) levels (observed in Fig. 4.3E) were not correlated with increased Rheb-GTP activity. Garami *et al.* also measured the levels of GTP bound Rheb in both wild-type and TSC2 knockout cells. Because of TSC2's regulatory role, a TSC2 knockout significantly increased Rheb-GTP levels ($\epsilon_{pred} = 0.10 \pm 0.03$, $\epsilon_{rand} = 0.09 \pm 0.06$) (Fig. 4.4C). Lastly, the model predicted the levels of 4E-BP1 bound eIF4E in response to heat shock

Table 4.2: Blind Prediction (P#) list along with species, cell type, prediction error, random error with a randomly generated parameter set and the corresponding literature reference.

P#	Species	Cell Type	Compartment	Prediction	Random	Source
P1	43S-mRNA (GTP)	rabbit reticulocytes	in vitro	0.52 ± 0.40	0.82 ± 0.51	[316]
P2	Rheb-GTP	A14 NIH 3T3 cells	Total lysate	0.22 ± 0.11	0.42 ± 0.01	[151]
P3	Rheb-GTP	A14 NIH 3T3 cells	Total lysate	0.10 ± 0.03	0.09 ± 0.06	[151]
P4	eIF4E:4EBP1	CHO K1 cells	Total lysate	0.51 ± 0.33	1.67 ± 1.17	[570]
P5	pAkt(Ser473)	HEK293 cells	Total lysate	0.27 ± 0.09	0.72 ± 0.09	[151]

(ϵ_{pred} = 0.51 ± 0.33, ϵ_{rand} = 1.67 ± 1.17) (Fig. 4.4D) [570]. Because the model was not trained on stress-induced translation inhibition, this result further demonstrated the predictive power of the model population.

4.3.3 Sensitivity analysis identified robust and fragile features of the initiation architecture.

Sensitivity analysis generated falsifiable predictions about the fragility or robustness of structural features of the initiation architecture. First order sensitivity coefficients were computed for 40 parameter sets selected from the ensemble (materials and methods), time-averaged and rank-ordered for the 250 species in the model, in the presence and absence of insulin and IRS-1 feedback. The sensitive components of insulin signaling shifted from Rheb in the absence of insulin to a combination of Rheb and PI3K in the presence of insulin. Sensitivity coefficients (s_{ij}) were calculated with and without insulin over the complete 100 min response (Fig. 4.5A). Globally, processes involved with 80S formation were consistently ranked among the most sensitive, irrespective of insulin. However,

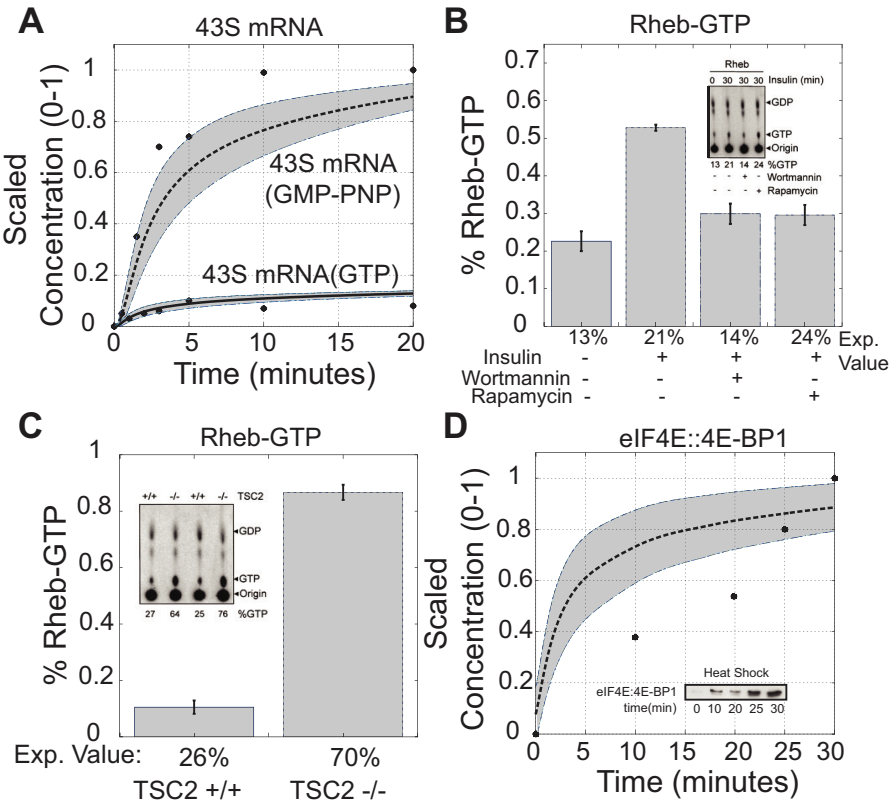


Figure 4.4: Blind model predictions for the ensemble (N = 400). The predictive ability of model ensemble was assessed by comparing model performance with novel experimental data. Dotted lines represent the simulation mean of the ensemble, while the shaded region denotes the 99.9% confidence estimate for the mean.The solid dots represent the scaled experimental data. **A.** In vitro time course for formation of 43S-mRNA complex. A slowly-hydrolyzable GTP homologue (GMP-PNP) was used in place of GTP to isolate formation of this intermediate complex. GMP-PNP data was used for training while GTP data was used for validation. **B.** Percent of Rheb-GTP to Rheb-GDP in the presence of insulin, wortmannin and rapamycin (A14 NIH 3T3 cells). **C.** Percent of Rheb-GTP to Rheb-GDP in wildtype and TSC2 lacking cells (MEF cells). **D.** 4EBP1 bound EIF4E in the presence of heat shock (CHO.K1 cells).

the sensitivity of other signal processing components changed with insulin status. For example, without insulin, Rheb/Rheb-GDP were highly fragile (rank ≥ 0.25), while PI3K, PIP2, PIP3 and PTEN were highly robust (rank ~ 0.0). Surprisingly, the relative sensitivity of these network components changed in the presence of insulin. While the fragility of Rheb/Rheb-GDP shifted modestly upward with insulin, the sensitivity of PI3K and its downstream complexes increased dramatically (rank ≥ 0.45) following insulin stimulation. This suggested that the combination of PI3K and Rheb activity was critical to insulin action over the full 100 min time window. However, it was unclear whether PI3K was always important, or if there was a temporal window in which PI3K became important following insulin stimulation. To explore this question, we time-averaged the sensitivity coefficients over early- and late-phase time periods following insulin stimulation (Fig. 4.5B). The 0-5 minute time period captured the initial network dynamics, while the 30-100 minute time period captured the network at a quasi-steady state. Generally, network components were more sensitive under dynamic operation (species beneath the 45^o line), compared with steady state. However, there were exceptions to this trend. For example, PI3K, PTEN and TSC1/2 were equally sensitive in both time frames, suggesting these species played important roles in both dynamic and steady state signaling. On the other hand, the Rheb rank decreased from 0.6 to 0.25 as the network moved toward steady state. Taken together, the sensitivity results suggested that Rheb activity controlled the background level of translation initiation while the PI3K axis in combination with Rheb regulated insulin-induced initiation. Moreover, the transition between PTEN and PI3K control occurred directly after the addition of insulin, giving rise to switch like behavior.

IRS-1 phosphorylation, a well known negative feedback mechanism [322,

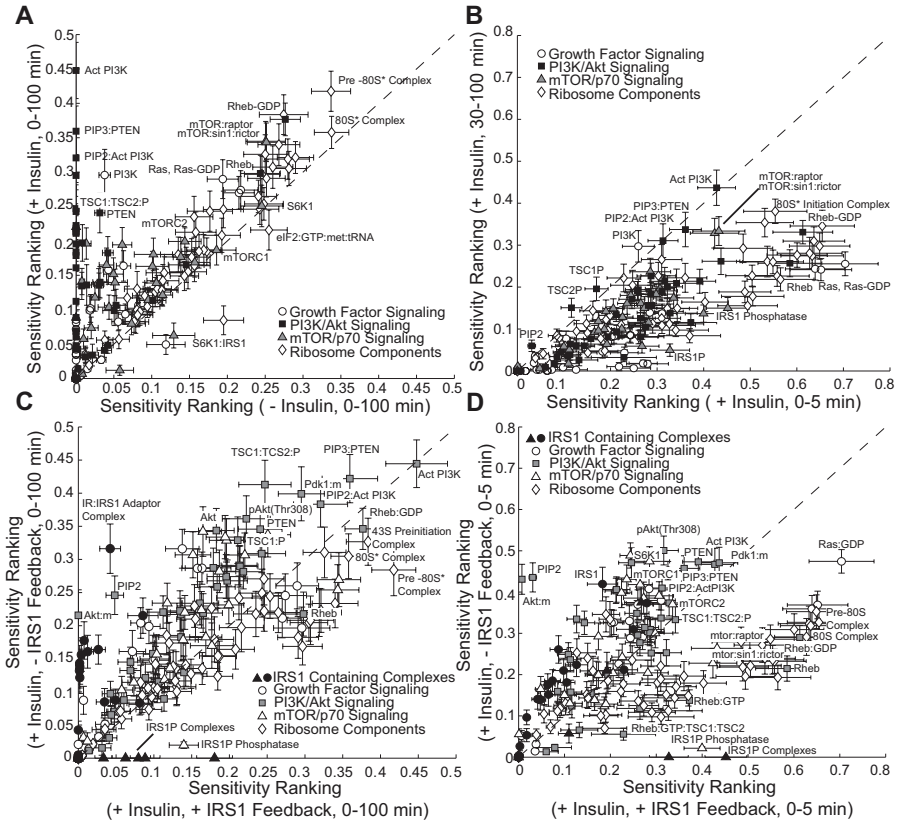


Figure 4.5: Sensitivity analysis of a population of initiation models (N = 40). Species with a high sensitivity ranking are considered fragile while species with a low sensitivity ranking are considered robust. **A.** Sensitivity ranking of network species in the presence and absence of insulin. **B.** Time-course sensitivity ranking of network species. **C,D.** Sensitivity ranking of network species in the presence and absence of IRS-1 feedback. Black fill denoted complexes containing IRS-1, grey fill denotes PI3K/Akt associated signaling components. Sensitivity values were time averaged over 0-100 minutes and 0-5 minutes, respectively. Error bars denote one standard error in the sensitivity ranking computed over a family of uncorrelated (mean correlation of approximately 0.6) parameter sets selected for the analysis.

551, 550, 111], attenuated PI3K sensitivity. We explored the role of IRS-1 feedback by comparing sensitivity coefficients under insulin stimulation in the presence and absence of IRS-1 feedback (Fig. 4.5C). The most significant change without feedback was the sensitivity of the IR:IRS-1 and adaptor complexes (Fig. 4.5C, black fill); IR:IRS-1, which anchors the adaptor complex to the activated receptor and is immediately upstream of PI3K activation, changed from NSS rank $\simeq$ 0.04 to 0.32. The sensitivity of the PI3K/Akt signaling axis also increased in the absence of feedback (Fig. 4.5C, grey fill). Surprisingly, the sensitivity of Rheb and many ribosomal components decreased in the absence of feedback. Similar results were observed when sensitivity coefficients were time averaged over the 0 to 5 min time window (Fig. 4.5D). These sensitivity calculations suggest that IRS-1 feedback plays a significant role in insulin signaling by modulating the relative importance of PI3K versus Rheb. Thus, IRS-1 feedback though not directly identified as a fragile regulatory motif, has significant effects on network function.

Lastly, the architectural features of the initiation network identified by sensitivity analysis, as either fragile or robust, were likely parameter independent. While first-order sensitivity coefficients are local, we sampled a family of uncorrelated parameter sets (mean correlation of approximately 0.6) to generate a set of consensus conclusions. By sampling over many uncorrelated sets, we calculated how our conclusions changed with different unrelated parameter sets. The distribution of ranking (standard-error shown in Fig. 4.5) suggested that despite parametric uncertainty, sensitivity analysis over an uncorrelated model population produced a consensus estimate of the strongly fragile or robust elements of the insulin signaling network. Previously, we (and others) have shown that monte-carlo parameter set sampling produced similar results in several studies

across many signaling networks [510, 320, 74, 527].

4.3.4 Robustness analysis identified key regulators of translation initiation

Knockdown simulations were conducted for 92 proteins to estimate the functional connectedness of the initiation network. The effects of the perturbations were quantified by calculating the relative change (α) in translational activity (80S formation) for each simulated knockout in the presence (Fig. 4.6A) and absence (Fig. 4.6B) of insulin. Knockdown simulations were conducted using 400 models selected from the ensemble based on error and correlation (materials and methods). Proteins were classified based on their impact on translational activity: little or no effect ($\alpha \approx 1$, white fill), moderate decrease ($\alpha \approx 0.6$, dark grey), critical ($\alpha \approx 0$, light grey) and increase ($\alpha > 1$, black). Generally, knockdowns in the presence of insulin were more likely to decrease initiation (Fig. 4.6A). Knockdown analysis identified 24 proteins (or 26% of the network) that were critical to translation initiation irrespective of insulin status; these critical components included mTORC1, S6K1, several initiation factors and other ribosomal components. Sensitivity analysis suggested basal translation was governed by Rheb, while insulin-induced initiation was governed by PI3K. Robustness analysis showed that perturbations in PI3K signaling, in the presence of insulin, restored initiation control to Rheb. Initiation was reduced by 40% by disrupting species immediately upstream or downstream of PI3K; a moderate reduction in the presence of insulin demonstrated that initiation was governed by both PI3K and Rheb. Lastly, deletion of TSC1/2 (negative regulator of Rheb)

or 4E-BP1 (sequesters the cap-binding protein eIF4E), increased initiation in the presence of insulin. Interestingly, for several proteins the direction or magnitude of change in initiation activity depended upon the presence or absence of insulin. For example, PTEN deletion significantly increased initiation ($\alpha \gg 1$) in the absence of insulin, but had no effect when insulin was present. On the other hand, PI3K deletion had a moderate reduction on 80S formation in the presence of insulin, but only a small effect in the absence of insulin (Fig. 4.6B). These results suggested that PI3K and PTEN were conditionally fragile proteins; in the presence of insulin, PI3K is a critical signal processing node, while PTEN acts to restrain inadvertent basal initiation.

Paradoxically, Rheb and mTORC2 subunit (sin1, rictor) knockdowns increased initiation. Our expectation from sensitivity analysis was that a Rheb knockdown would reduce initiation, irrespective of insulin status. However, this was not universally true; some members of the model population showed increased initiation (Fig. 4.6C). Following the deletion of PTEN, approximately 80% (or 323 of the 400 models sampled) had increased initiation in the absence of insulin. Of these models, 16% (or 51 of 323) had at least a two fold increase in translational activity. This result was expected; deletion of a protein species resulted in a qualitatively similar change in initiation across the ensemble of models. However, for Rheb knockdowns, members of the ensemble demonstrated qualitatively different behavior. For 84% (or 334 of 400) of the models sampled, Rheb knockdowns significantly down-regulated initiation. Thus, the vast majority of models behaved as expected. Interestingly, 20 models (or 5% of the models sampled) had increased translation initiation in the presence of a Rheb knockdown, with 15 models demonstrating greater than a two-fold change (Fig. 4.6C). Thus, the model population estimated by POETs contained

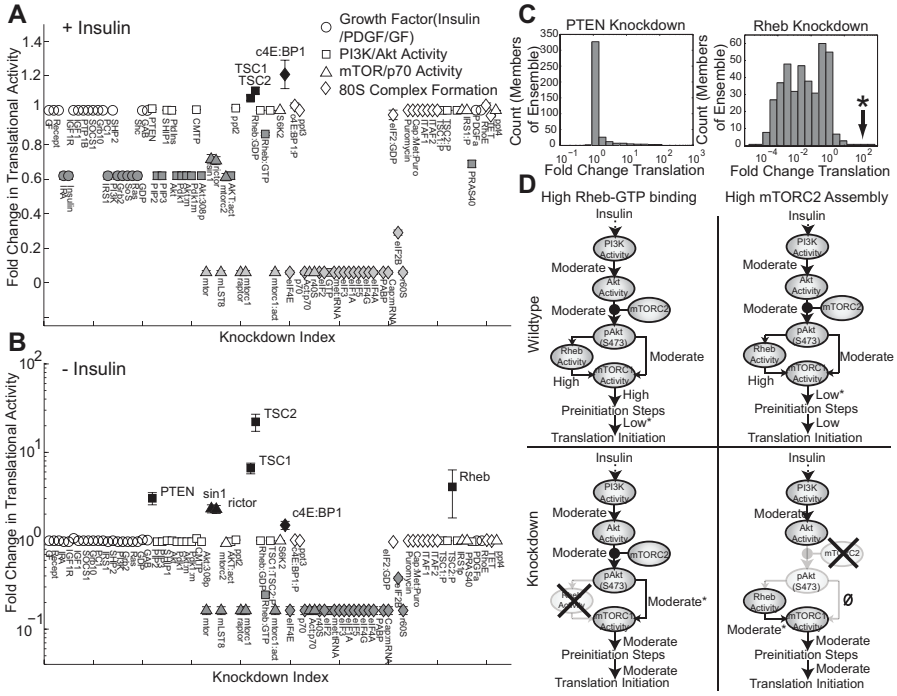


Figure 4.6: Species knockdown simulations for a population of translation initiation models (N = 400). Simulated knockdowns were performed by removing nodes from the stoichiometric matrix. The relative change in 80S formation resulting from the removal of a species was used to quantify the impact of the knockdown. **A.** Species knockdowns in the presence of insulin. Simulated knockdowns resulted in increased (black), constant (white), moderately decreased (dark grey) or severely decreased (light grey) translational levels. **B.** Species knockouts in the absence of insulin. Simulated knockdowns resulted in increased (black), constant (white), or decreased (grey) translational levels. **C.** Histogram of translation levels across each member of parameter ensemble. Asterisk index indicates parameter sets that were selected for further analysis. **D.** Alternative modes of network operation. For a subset of the ensemble, initiation increased following Rheb or mTORC2 disruption. Asterisk indicates rate-limiting step.

models with qualitatively different behavior. Histograms of sin1 and rictor knockdowns showed a similar trend (results not shown). We explored the flux vectors of these outlying parameter sets to better understand the mechanistic effect of Rheb and rictor/sin1 knockouts. All of the outlying models were in regions of parameter space where the association between Rheb and GTP was very high. Strong Rheb/GTP binding resulted in abnormally high signal flux to mTORC1 despite the inhibitory effects of TSC1/2 (Fig. 4.6D, top-left). Consequently, less GTP was available for the energy-dependent steps of translation initiation (i.e. formation of eIF2-GTP-met-tRNA tertiary complex). Additionally, strong association between Rheb and GTP resulted in high levels of activated mTORC1 and S6K1. However, despite the high levels of mTORC1, GTP-dependent pre-initiation reactions were rate limiting (Fig. 4.6D, labeled*). Thus, Rheb knockdown released the network from its GTP limitation and shifted the predominant signaling mode to mTORC2. This shift in signaling, while lowering the activated mTORC1/S6K1 level, ultimately resulted in higher levels of initiation (Fig. 4.6 bottom-left). On the other hand, the rictor/sin1 knockdown behaved differently. The rate-limiting step for the rictor/sin1 knockdowns was mTORC1 activation: more Rheb-GTP was present than there was mTORC1 to be activated (Fig. 4.6D top-right). Thus, knockdown of rictor/sin1 prevented the assembly of mTORC2 and freed the mTOR subunit to be used for mTORC1 assembly. This shift toward mTORC1 assembly and activation relieved the Rheb-GTP/mTORC1 bottleneck, resulting in increased initiation.

4.4 Discussion

In this study, we developed and analyzed a population of insulin and growth factor induced translation initiation models. These models described the integration of insulin and growth-factor signals with 80S assembly. A family of model parameters was estimated from 24 transient and steady state data sets using multiobjective optimization. In addition to the training data, the model family also predicted novel data sets not used during model training. The population of initiation models was analyzed using sensitivity and robustness analysis to identify the key components of insulin-induced translation initiation. Without insulin, a balance between the pro-initiation activity of the GTP-binding protein Rheb and anti-initiation activity of PTEN controlled basal initiation. Rheb knockdown simulations confirmed decreased initiation in the majority of the model population. Surprisingly, we also identified a model subpopulation in which deletion of Rheb or mTORC2 components increased initiation. In these cases, removal of Rheb or mTORC2 components relieved a rate-limiting bottleneck e.g., constrained levels of GTP, leading to increased initiation. On the other hand, in the absence of insulin, translation initiation increased for all models in the population following a PTEN deletion. In the presence of insulin, Rheb and PTEN were no longer the dominant arbiters of initiation; a combination of PI3K and Rheb activity controlled inducible initiation, where PI3K was only critical in the presence of insulin. PI3K deletion in the presence of insulin removed the ability of the network to process insulin signals, but did not remove initiation altogether. PI3K deletion reduced initiation to approximately 60% of its maximum level. Interestingly, the relative contribution of PI3K versus Rheb to the overall initiation level could be tuned by IRS-1 feedback. In the absence of

feedback, PI3K was more important than Rheb to signal propagation, while the opposite was true in the presence of feedback.

PI3K and PTEN in combination with Rheb are components of a switch that regulates inducible and basal translation initiation. In the absence of insulin, a balance between the pro-initiation activity of Rheb and the anti-initiation activity of PTEN regulated basal initiation. On the other hand, in the presence of insulin, control shifted to a combination of Rheb and PI3K, where PI3K activity regulated the inducible fraction of initiation. Thus, deletion of PTEN, constitutive activation of PI3K or constitutively active Rheb could all induce aberrant translation initiation without an insulin or growth factor signal. Yuan and Cantley noted that every major species in the PI3K pathway is mutated or over-expressed in a wide variety of solid tumors [619]. For example, activating mutations in PIK3CA, the gene encoding the catalytic subunit of PI3K, induces oncogene signaling in colon, brain and gastric cancers [468]. On the other hand, PTEN mutations have long been implicated in a spectrum of cancer types [465]. Both PIK3CA and PTEN mutations induce a pro-initiation operational mode in the absence of growth factor. Likewise, constitutive Rheb activity induces a variety of pleiotropic traits involving translation. For example, Saucedo *et al.* showed that Rheb over-expression in *Drosophila melanogaster* increased cell size, wing area and G1/S cell cycle progression [473]. Rheb and TSC1/2 mutations are also frequently observed in cancer [235, 311]. Taken together, our study supports the supposition of Taniguchi *et al.* that PI3K is a critical arbiter of insulin-induced translation initiation [525]. However, we have also shown that initiation control and particularly the role of PI3K was more nuanced; while insulin or growth-factor inducible initiation was controlled by PI3K, basal initiation was controlled by Rheb. Moreover, in the absence of insulin, PTEN was

the critical upstream initiation regulator, not PI3K. This suggested that the relative level of the phosphorylated phospholipids PIP2 and PIP3 was actually the key mediator of initiation. Lastly, Taniguchi *et al.* suggested that Akt was also a key node involved in insulin action. Our previous model directly supports this, however, the current model does not. Rather, our analysis suggested that Rheb was the downstream controller of initiation. These two points of view are not contradictory however, as Rheb activation is driven by phosphorylated Akt.

The initiation model connectivity was assembled from an extensive literature review, however, several potentially important signaling mechanisms were not included. First, we should revisit the role of PRAS40. Currently, PRAS40 acts as a cofactor that aids in pAkt(Ser473)-mediated activation of mTORC1. Sancak *et al* suggested that PRAS40 sequesters mTORC1, and only after phosphorylation by Akt does it releases from mTORC1 [469]. Other groups have also shown that mTORC1 can phosphorylate and inhibit PRAS40, thus providing a positive feedback mechanism for Akt-mediated mTORC1 activation [132, 572]. A more complete description of PRAS40 will enhance our ability to interrogate Akt dependent mTORC1 activation. Second, we need to refine the description of IRS-1 feedback. Currently, we assume a single deactivating phosphorylation event at Ser308. However, several studies have shown that IRS-1 can be phosphorylated at multiple serine sites, which are both activating and deactivating [414, 173]. Additionally, PTEN is known to dephosphorylate activated PDGF receptors and attenuate their activity, a feature not included currently [519]. A more complete description of IRS-1 phosphorylation could help define how, and under what conditions, IRS-1 regulation attenuates PI3K activation. Third, we modeled the regulation of 4E-BPx as a single phosphorylation event where phosphorylated 4E-BPx was unable to bind to eIF4E. In

reality, 4E-BPx family members, such as 4E-BP1, have several phosphorylation sites [442] and the release of eIF4E is driven only after multiple conserved phosphorylation events [162]. Additionally, eIF4E can itself be phosphorylated at Ser209; while there is agreement that the phosphorylation of eIF4E does have a regulatory significance, the data is contradictory as to whether it is positive or negative [474]. Fourth, signaling downstream of mTORC1 has also been shown to mediate translation modes beyond those included in our model. eIF3 has been identified as a scaffolding protein that recruits mTORC1 to untranslated mRNA and facilitates S6K1 and 4E-BP1 phosphorylation [214]. S6K1 can also activate eIF4B, a protein that helps eIF4A to unwind the secondary structure of untranslated mRNA [443]. Further, a recently discovered scaffold protein, SKAR, has been shown to assist S6K1 recruitment to mRNA [323]. Lastly, because of mTORC1's unique cellular role, it would be interesting to explore how other aspects of metabolism interact with insulin signaling to mediate decisions between translation, lipid synthesis or proliferation. In these studies, one could imagine constructing *in-vivo* mouse models to explore the physiological role of mTORC1 signaling in important diseases such as diabetes or cancer.

4.5 Materials and Methods

4.5.1 Formulation and solution of the model equations.

The translation initiation model was formulated as a set of coupled non-linear ordinary differential equations (ODEs):

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{r}(\mathbf{x}, \mathbf{p}) \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{4.1}$$

The symbol $\mathbf{S}$ denotes the stoichiometric matrix (250×573). The quantity $\mathbf{x}$ denotes the concentration vector of proteins (250×1). The term $\mathbf{r}(\mathbf{x}, \mathbf{p})$ denotes the vector of reaction rates (573×1). The (i, j) element of the matrix $\mathbf{S}$, denoted by σ_{ij}, described how protein i was involved in rate j. If $\sigma_{ij} < 0$, then protein i was consumed in r_j. Conversely, if $\sigma_{ij} > 0$, protein i was produced by r_j. Lastly, if $\sigma_{ij} = 0$, then protein i was not involved in rate j. We assumed mass-action kinetics for each interaction in the network. The rate expression for interaction q was given by:

$$r_q\left(\mathbf{x}, k_q\right) = k_q \prod_{j \in \{\mathbf{R}_q\}} x_j^{-\sigma_{jq}} \tag{4.2}$$

The set $\left\{\mathbf{R}_q\right\}$ denotes reactants for reaction q while σ_{jq} denotes the stoichiometric coefficient (element of the matrix $\mathbf{S}$) governing species j in reaction q. The quantity k_q denotes the rate constant governing reaction q. All reversible interactions were split into two irreversible steps. Model equations were generated using UNIVERSAL from an SBML input file (available in the supplemental materials). UNIVERSAL is an open source Objective-C/Java code generator, which is freely available as a Google Code project (http://code.google.com/p/universal-code-generator/). The model equations were solved using the LSODE routine in OCTAVE (v 3.0.5; www.octave.org) on an Apple workstation (Apple, Cupertino, CA; OS X v10.6.4).

When calculating the response of the model to the addition of insulin or other growth factors, we first ran to steady state and then issued the perturbation. The steady state was estimated numerically by repeatedly solving the model equations and estimating the difference between subsequent time points:

$$\|\mathbf{x}(t + \Delta t) - \mathbf{x}(t)\|_2 \leq \gamma \tag{4.3}$$

The quantities $\mathbf{x}(t)$ and $\mathbf{x}(t + \Delta t)$ denote the simulated concentration vector at

time t and $t + \Delta t$, respectively. The L_2 vector-norm was used as the distance metric, where $\Delta t = 1$ s and $\gamma = 0.001$ for all simulations.

4.5.2 Estimation and cross-validation of a population of models using Pareto Optimal Ensemble Techniques (POETs).

We used multiobjective optimization in combination with cross-validation to estimate an ensemble of initiation models. Multiobjective optimization in combination with cross-validation allowed us to address qualitative conflicts in the training data, and to protect against model over-training. While computationally more complex than single-objective formulations, multiobjective optimization is an important tool to address qualitative conflicts in training data that arise from experimental error or cell-line artifacts [182]. Multiobjective optimization balances these conflicts allowing us to identify a consensus model population. In this study we used the Pareto Optimal Ensemble Technique (POETs) to perform the optimization. POETs integrates standard search strategies e.g., Simulated Annealing (SA) or Pattern Search (PS) with a Pareto-rank fitness assignment [504]. Denote a candidate parameter set at iteration $i + 1$ as $\mathbf{k}_{i+1}$. The squared error for $\mathbf{k}_{i+1}$ for training set j was defined as:

$$E_j(\mathbf{k}) = \sum_{i=1}^{\mathcal{T}_j} \left(\hat{\mathcal{M}}_{ij} - \hat{y}_{ij}(\mathbf{k}) \right)^2 \tag{4.4}$$

The symbol $\hat{\mathcal{M}}_{ij}$ denotes scaled experimental observations (from training set j) while $\hat{y}_{ij}$ denotes the scaled simulation output (from training set j). The quantity i denotes the sampled time-index and $\mathcal{T}_j$ denotes the number of time points for experiment j. In this study, the experimental data used for model training was typically the band intensity from immunoblots, where intensity was estimated

using the ImageJ software package [1]. The scaled measurement for species x at time $i = \{t_1, t_2, .., t_n\}$ in condition j is given by:

$$\hat{\mathcal{M}}_{ij} = \frac{\mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}}{\max_i \mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}} \tag{4.5}$$

Under this scaling, the lowest intensity band equaled zero while the highest intensity band equaled one. A similar scaling was defined for the simulation output. By doing this scaling, we trained the model on the relative change in blot intensity, over conditions or time (depending upon the experiment). Thus, when using multiple data sets (possibly from different sources) that were qualitatively similar but quantitatively different e.g., slightly different blot intensities over time or condition, we captured the underlying trends in the scaled data.

We computed the Pareto rank of $\mathbf{k}_{i+1}$ by comparing the simulation error at iteration $i+1$ against the simulation archive $\mathbf{K}_i$. We used the Fonseca and Fleming ranking scheme [133] to estimate the number of parameter sets that dominate $\mathbf{k}_{i+1}$. Parameter sets with increasing rank are progressively further away from the optimal trade-off surface. The parameter set $\mathbf{k}_{i+1}$ was accepted or rejected by POETs with probability $\mathcal{P}(\mathbf{k}_{i+1})$:

$$\mathcal{P}(\mathbf{k}_{i+1}) \equiv \exp\{-rank(\mathbf{k}_{i+1} \mid \mathbf{K}_i)/T\} \tag{4.6}$$

where T is the annealing temperature and $rank(\mathbf{k}_{i+1} \mid \mathbf{K}_i)$ denotes the Pareto rank for $\mathbf{k}_{i+1}$. The annealing temperature was discretized into 10 quanta between T_o and T_f and adjusted according to the schedule $T_k = \beta^k T_0$ where β was defined as $\left(T_f/T_o\right)^{1/10}$. The initial temperature was given by $T_o = n/log(2)$, where $n = 4$ was used in this study and the final temperature was $T_f = 0.1$. The epoch-counter k was incremented after the addition of 100 members to the ensemble. Thus, as the ensemble grew, the likelihood of accepting parameter sets with a large Pareto rank decreased. To generate parameter diversity, we randomly

perturbed each parameter by $\leq \pm 25\%$. We performed a local pattern search every q steps to minimize the residual for a single randomly selected objective. The local pattern-search algorithm has been described previously [149].

A leave-three-out cross-validation strategy was used to simultaneously calculate the training and prediction error during the parameter estimation procedure [261]. The 24 training data sets were partitioned into eight subsets, each containing 21 data sets for training and three data sets for validation. The leave-three-out scheme generated 18,886 probable models. From the approximately 6000 rank zero models, we iteratively selected 50 random models from each cross-validation trial with the lowest correlation and shortest Euclidian distance to the origin (minimum error). This selection technique produced sub-ensembles with low set-to-set correlation (≤ 0.50) and minimum training error.

4.5.3 Sensitivity and robustness analysis of the initiation model population.

Sensitivity coefficients were calculated for 40 models selected from the ensemble (rank-zero, low-correlation, minimum error selection). First-order sensitivity coefficients at time t_q:

$$s_{ij}\left(t_q\right) = \left.\frac{\partial x_i}{\partial k_j}\right|_{t_q} \tag{4.7}$$

were computed by solving the kinetic-sensitivity equations [105]:

$$\begin{pmatrix} d\mathbf{x}/dt \\ d\mathbf{s}_j/dt \end{pmatrix} = \begin{bmatrix} \mathbf{S}\cdot\mathbf{r}\left(\mathbf{x},\mathbf{k}\right) \\ \mathbf{A}\left(t\right)\mathbf{s}_j + \mathbf{b}_j\left(t\right) \end{bmatrix} \qquad j = 1, 2, \ldots, P \tag{4.8}$$

subject to the initial condition $\mathbf{s}_j(t_0) = \mathbf{0}$. The quantity j denotes the param-

eter index, P denotes the number of parameters in the model, $\mathbf{A}$ denotes the Jacobian matrix, and $\mathbf{b}_j$ denotes the jth column of the matrix of first-derivatives of the mass balances with respect to the parameters. Sensitivity coefficients were calculated by repeatedly solving the extended kinetic-sensitivity system for forty parameters sets selected from the final 400 member ensemble. These sets were chosen to be comparable to the final 400 member ensemble on the basis of parametric coefficient of variation (CV); the sets selected for sensitivity analysis had a mean CV of 0.85 ± 0.5 and a mean correlation of approximately 0.6. Thus, there were diverse and uncorrelated. The Jacobian $\mathbf{A}$ and the $\mathbf{b}_j$ vector were calculated at each time step using their analytical expressions generated by UNIVERSAL.

The resulting sensitivity coefficients were scaled and time-averaged (Trapezoid rule):

$$\mathcal{N}_{ij} \equiv \frac{1}{T} \int_0^T dt \cdot |s_{ij}(t)| \tag{4.9}$$

where T denotes the final simulation time. The time-averaged sensitivity coefficients were then organized into an array for each ensemble member:

$$\mathcal{N}^{(\epsilon)} = \begin{pmatrix} \mathcal{N}_{11}^{(\epsilon)} & \mathcal{N}_{12}^{(\epsilon)} & \dots & \mathcal{N}_{1j}^{(\epsilon)} & \dots & \mathcal{N}_{1P}^{(\epsilon)} \\ \mathcal{N}_{21}^{(\epsilon)} & \mathcal{N}_{22}^{(\epsilon)} & \dots & \mathcal{N}_{2j}^{(\epsilon)} & \dots & \mathcal{N}_{2P}^{(\epsilon)} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathcal{N}_{M1}^{(\epsilon)} & \mathcal{N}_{M2}^{(\epsilon)} & \dots & \mathcal{N}_{Mj}^{(\epsilon)} & \dots & \mathcal{N}_{MP}^{(\epsilon)} \end{pmatrix} \qquad \epsilon = 1, 2, \dots, N_\epsilon \tag{4.10}$$

where ϵ denotes the index of the ensemble member, P denotes the number of parameters, N_ϵ denotes the number of ensemble samples and M denotes the number of model species. To estimate the relative fragility or robustness of species and reactions in the network, we decomposed the $\mathcal{N}^{(\epsilon)}$ matrix using Singular Value Decomposition (SVD):

$$\mathcal{N}^{(\epsilon)} = \mathbf{U}^{(\epsilon)} \mathbf{\Sigma}^{(\epsilon)} \mathbf{V}^{T,(\epsilon)} \tag{4.11}$$

Coefficients of the left (right) singular vectors corresponding to largest $\theta \leq 15$ singular values of $\mathcal{N}^{(\epsilon)}$ were rank-ordered to estimate important species (reaction) combinations. Only coefficients with magnitude greater than a threshold ($\delta = 0.001$) were considered. The fraction of the θ vectors in which a reaction or species index occurred was used to determine its importance (sensitivity ranking). The sensitivity ranking was compared between different conditions to understand how control in the network shifted as a function of perturbation or time (Fig. 4.5).

Robustness coefficients were calculated as shown previously [527]. Robustness coefficients (denoted by $\alpha\left(i, j, t_o, t_f\right)$) are the ratio of the integrated concentration of a network marker in the presence (numerator) and absence (denominator) of a structural or operational perturbation. The quantities t_0 and t_f denote the initial and final simulation time, respectively, while i and j denote the indices for the marker and the perturbation respectively. If $\alpha\left(i, j, t_o, t_f\right) > 1$, then the perturbation *increased* the marker concentration. Conversely, if $\alpha\left(i, j, t_o, t_f\right) < 1$ the perturbation *decreased* the marker concentration. Lastly, if $\alpha\left(i, j, t_o, t_f\right) \sim 1$ the perturbation did not influence the marker concentration. Robustness coefficients were calculated over 400 models selected from the ensemble (rank-zero, low-correlation, minimum error selection). Convergence analysis suggested that the qualitative conclusions drawn from the robustness analysis would not change if more than N = 400 parameter sets were sampled (Fig. 4.7).

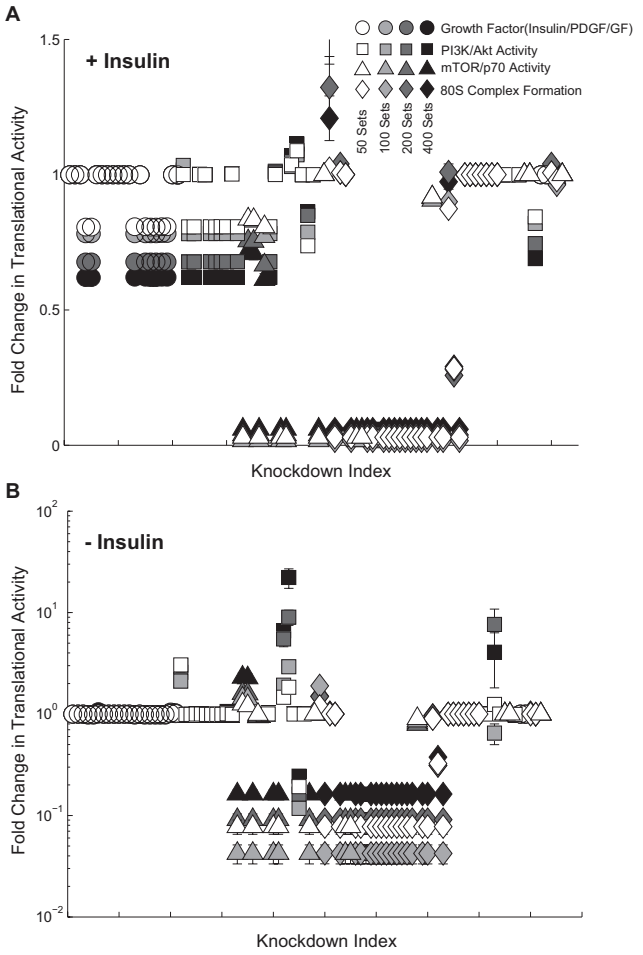


Figure 4.7: Effect of the ensemble size on the knockdown simulations. Fold change of the translational activity was calculated for ensemble sizes of N = 50 (white fill), N = 100 (light grey), N = 200 (dark grey) and N= 400 (black) randomly selected parameter sets in the presence and absence of insulin. For the majority of the perturbations, the robustness coefficients converged for as few as 50 parameter sets. In a small number of other cases, the robustness coefficients varied significantly up to 200 parameter sets. Between 200 - 400 sets the robustness coefficients largely converged to qualitatively and quantitatively similar answers.

CHAPTER 5

ANALYSIS OF THE CORE ARCHITECTURE REGULATING TGFβ INDUCED EMT

Authors - Gould R, Chakrabarti A, Butcher JB and Varner JD.

Submitted to - Mol. Sys. Biol.

5.1 Abstract

Transforming growth factor β (TGFβ) has been shown as a potent inducer of epithelial to mesenchymal transition (EMT) in both embryonic and pathological conditions. The complexity of TGFβ signaling is overwhelming due to the large numbers of interacting protein complexes, complicated feedback mechanisms, and cross-talk between multiple signaling pathways. As a result, even understanding the fundamental regulation of cell specific markers remains difficult. Systems biology has been suggested as an essential tool for understanding the orchestration of EMT and omission of such a critical contributor may lead to an inaccurate understanding. Thus, we modeled the molecular interactions of TGFβ induced EMT using mass action kinetics within an ordinary differential equation (ODE) based framework. 2695 unknown model parameters (1700 kinetic constants and 995 non-zero initial conditions) were estimated using 41 steady-state experimental data sets taken from literature sources. Using POETs we implemented a population based approach to identify different operational paradigms within EMT. Using signal flow, sensitivity, and robustness analysis, our model suggested three important levels of regulation. The differential role of AP1/SP1, selective de-phosphorylation of the MAPK/ERK cascade, and the

availability of LEF1. These results provide insight into the core molecular architecture of TGFβ induced EMT and reveal possible transformational aspects between cellular phenotype.

5.2 Introduction

Epithelial to mesenchymal transition (EMT) is an important phenomenon initiating the development of tissues and other morphogenetic events. This process is associated with a loss of epithelial traits, such as E-cadherin, and the acquisition of mesenchymal characteristics, such as Vimentin, and Fibronectin [520, 521]. TGFβ has been shown an important inducer of this process as seen by the initiating the formation of endocardial cushions in the atrioventricular valves and mediating palate fusion of medial epithelial cells [620]. In this context, EMT terminates and has been suggested as a highly controlled process [290]. While being an integral process during development, EMT is also recapitulated under pathological conditions, prominently in fibrosis and in invasion and metastasis of carcinomas [533]. The complexity of TGFβ signaling is overwhelming due to the large numbers of interacting protein complexes, complicated feedback mechanisms, and cross-talk between multiple signaling pathways [534]. As a result, even understanding the regulation of cell specific markers remains difficult.

Systems biology has been suggested as an essential tool for understanding the dynamic integration of such complex pathways and omission of such a critical contributor may lead to inaccurate understanding of EMT [3]. Current computation models investigating TGFβ signaling have proven insightful. For ex-

ample, Chung et al. constructed an ordinary differential equation model using mass action kinetics to investigate the mechanisms associated with receptor activation and Smad signaling. Their model suggested that a reduction in the level of functional TGFβ receptors in cancer cells may lead to attenuated and slower TGFβ stimulated signaling responses via Smad2 [76]. Similar work by Vilar et al. suggested specific changes in receptor trafficking patterns that can lead to phenotypes that favor tumor progression [565]. Although these models have provided useful information about receptor dynamics, they are limited in complexity, and understanding global phenomenons such as TGFβ induced EMT cannot be considered.

Investigating large scale networks have proven to be difficult. This is due to the structural/parametric uncertainty associated with all deterministic models and lack of computational approaches capable of handling such vast systems. Recent methods to minimize parametric uncertainty have suggested a population based approach rather than a single best-fit but uncertain model. While, ensembles often constrain individual parameter values poorly, they can still robustly capture model predictions [176]. Tasseff et al. used this approach to effectively capture which signaling systems were important in the loss of androgen dependence [526]. Furthermore, the application of advanced sampling techniques have demonstrated the capability to investigate extensive parameter space. For example, multi-objective optimization ensemble techniques can be used to balance conflicts in the training data [505].

Recent work by Medici et al. has revealed key molecular pathways through which sequential activation of TGFβ isoforms induces EMT. This framework culminates decades of work within the EMT field and has provided a structural

basis through which systems biology can now be applied. Our objective was to develop a computational model which can quantitatively investigate the induction of EMT through MAPK and Smad dependent pathways. Model infrastructure components included TGFβ2 ligand binding, TGFβRI,II,III transmembrane growth factor receptor phosphorylation, activation of the Ras/MAPK pathways, autocrine response of the TGFβ3 ligand, activation of the Smad pathways, cytosolic release of β-catenin from E-cadherin, and complex cross-talk between the MAPK/Smad signaling pathways. We modeled the molecular interactions using mass action kinetics (describe the rate of each molecular interaction) within an ordinary differential equation (ODE) framework. Thus, 2695 unknown model parameters (1700 kinetic constants and 995 non-zero initial conditions) were estimated using 41 steady-state training sets taken from literature sources. Model parameters were estimated using experimental values conducted with DLD1 colon carcinoma, MDCKII, and A375 melanoma cells and taken from literature sources [347]. We implemented a population approach to estimate a family or ensemble of parameters that was consistent with the training data [346]. Using POETs, we extensively investigated parameter space to determine the best possible ensemble and cross-validate through a leave one out technique.

5.3 Results

5.3.1 Biological Background - EMT

TGFβ induced EMT is a complex process involving both the canonical MAPK and Smad pathways (Fig. 5.1). TGFβ2 signals through the RAS-RAF-MEK-ERK-AP1 pathway to up-regulate snail and slug expression [346]. Snail, a known repressor of junctional proteins, inhibits the expression of E-cadherin [61]. This initial repression of E-cadherin leads to a release of β-catenin from cell membrane. Cystolic β-catenin can then translocate to the nucleus and form transcriptional complexes with TCF-4 to drive TGFβ3 expression [347]. TGFβ3 signals to the cells interior by binding to type II receptors, which form heterodimers with type I receptors (ALK5, ALK2) [103]. This activates the receptors serine/threonine kinase activity to phosphorylate and activate the receptor Smads 2/3 [339]. To increase ligand affinity, β-glycan (TGFβR3) can also interact with TGFβRI,II which has been shown in both development and adult conditions [155]. Phosphorylated Smads 2/3 (pSmad2/3) then form heterodimers with partner Smad4 and translocate to the nucleus. pSmads complexes then up-regulate other transcription factors, such as LEF1 [347]. The pSmad2/4/LEF1 has been shown to directly repress the E-cadherin gene [367]. LEF1 also binds with β-catenin to upregulate mesenchymal proteins and further promote the mesenchymal to epithelial transition [347]. The MAPK and Smad pathways involve redundant feedback mechanisms involving both inhibitory and synergistic effects at the cytosolic and nuclear level [337]. In response to the MAPK pathway, Smad2/3 are phosphorylated by ERK at multiple MAP kinase sites in the linker region that negatively affect Smad function [338]. Downstream com-

ponents of MAP kinase signaling pathways, especially transcription factors of the AP1/SP1 family, interact with the R-Smad/ Smad4 complex, providing an additional level of crosstalk between these pathways [231].

5.3.2 Estimating the ensemble of TGFβ induced EMT models using POETs

EMT has been described as a complex process orchestrated by multiple signaling pathways, with numerous redundant feedback systems. Applying an ODE model to such a vast system makes it difficult to identify model parameters, even with extensive training data. However, population of model based approaches have been promising due to their ability to constrain model behavior despite uncertainty (structural and parametric). In this study, an ensemble of plausible model parameters was estimated for TGFβ induced cancer cell lines. Model connectivity (human) was assembled from literature and online databases; String-8 [233], NetworKIN [305] and TRANSFAC. The 2695 unknown model parameters (1700 kinetic constants and 995 non-zero initial conditions) were estimated using 41 steady-state training sets taken from literature sources [347]. Kinetic parameters and unspecified initial conditions were estimated using the POETs multiobjective optimization algorithm [504]. Additionally, we used cross-validation to independently estimate both the prediction and training error during the parameter estimation (Table 5.2). We used a leave-one-out strategy where one of the eleven objectives was reserved for validation while the remaining ten were used for model training. Thus, eleven different model families, trained and validated on different experimental data sets were

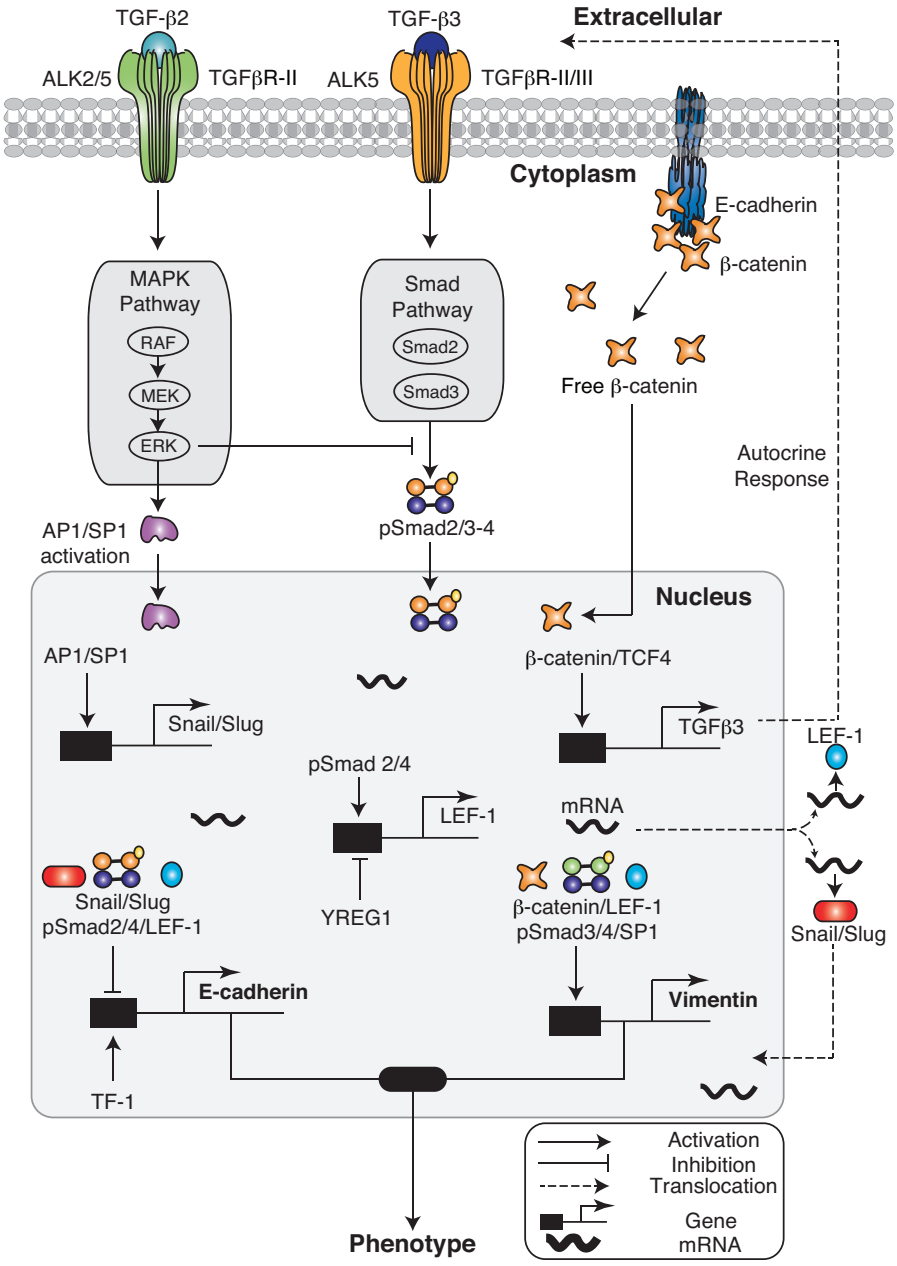


Figure 5.1: Schematic overview of the interaction network used in modeling the TGFβ induced EMT phenomenon. The model describes activation of the MAPK cascade through TGFβ2 followed by and autocrine response of TGFβ3 stimulating the Smad cascade. Cross-talk between MAPK and Smads has been shown at both the cytosolic and nuclear levels.

generated. POETs generated more than 15,000 probable models. We selected models with pareto rank one or less (N = 1091) for further study. The coefficient of variation (CV) for the parameters in the model ensemble ranged from 0.9 - 2.8. Approximately 50% of the model parameters were constrained with a CV $\leq$ 1.6. The most tightly constrained parameters governed LEF1 expression, transcription of Snail, and TGFβ3 regulation, while the least constrained parameters involved crosstalk between MAPK and transcriptional factor activation, e.g., ERK mediated AP1 activation. POETs was successful in identifying conflicts/relationships between the objective functions which were used for training/predictions.

Pareto fronts developed between several objectives, e.g., O3×O5, O7×O8, and O5×O8 (Fig. 5.3). These fronts indicated conflicts in the training data and/or an inability of the model to simultaneously capture different aspects of the training data. For example, O3 and O5 were measurements of the expression levels of TGFβ3 and the β-catenin/TCF4 complex through snail/slug induction. A front between these objectives suggested that overexpression of E-cadherin in O5 was conflicting with snail repression and therefore, not able to dissolve the β-catenin/TCF4 complex. On the other hand, a linear relationship developed between several objective functions, e.g., O2×O9 and O8×O11. These fronts indicated strong relationships between the training data and was effectively captured by the model. For example, O2 and O9 were measurements based upon the expression levels of Vimentin and E-cadherin through snail/slug and TGFβ3 induction (Fig. 5.3). The strong correlation between objectives suggested that the knockout of LEF1 was critical for maintaining E-cadherin expression in both cases.

OBJ-1	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
Snail	0.00	1.00	0.00			Snail	0.00	0.09	0.00			Snail	0.00	1.00	0.00			Snail	0.00	0.16	0.00		
Slug	0.00	0.00	1.00			Slug	0.00	0.00	0.19			Slug	0.00	0.00	1.00			Slug	0.00	0.00	0.19		
LEF1	0.00	0.73	1.00			LEF1	0.00	0.22	0.37			LEF1	0.00	0.66	1.00			LEF1	0.00	0.18	0.32		
Snail	-	-	+			Snail	-	-	+			Snail	-	-	+			Snail	-	-	+		
Slug	-	+	-			Slug	-	+	-			Slug	-	+	-			Slug	-	+	-		
OBJ-2	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
E-cad	0.92	0.00	0.00	1.00	0.81	E-cad	0.19	0.00	0.00	0.10	0.14	E-cad	1.00	0.00	0.00	0.06	0.06	E-cad	0.25	0.00	0.00	0.01	0.02
VIM	0.00	0.96	1.00	0.00	0.00	VIM	0.00	0.11	0.11	0.00	0.00	VIM	0.00	1.00	0.95	0.00	0.00	VIM	0.00	0.15	0.14	0.00	0.00
Snail	-	+	-	+	-	Snail	-	+	-	+	-	Snail	-	+	-	+	-	Snail	-	+	-	+	-
Slug	-	-	+	-	+	Slug	-	-	+	-	+	Slug	-	-	+	-	+	Slug	-	-	+	-	+
DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+
OBJ-3	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
TGFB3	0.00	1.00	0.95			LEF1	0.00	0.13	0.12			LEF1	0.00	0.77	1.00			LEF1	0.00	0.22	0.20		
Snail	-	-	+			Snail	-	-	+			Snail	-	-	+			Snail	-	-	+		
Slug	-	+	-			Slug	-	+	-			Slug	-	+	-			Slug	-	+	-		
OBJ-4	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
LEF1	0.00	0.98	1.00	0.03	0.02	LEF1	0.00	0.10	0.11	0.01	0.01	LEF1	0.00	0.95	1.00	0.00	0.00	LEF1	0.00	0.10	0.10	0.00	0.00
E-cad	1.00	0.00	0.00	0.54	0.35	E-cad	0.20	0.00	0.00	0.16	0.12	E-cad	1.00	0.34	0.37	0.00	0.37	E-cad	0.26	0.13	0.14	0.00	0.06
VIM	0.00	1.00	0.94	0.00	0.00	VIM	0.00	0.13	0.12	0.00	0.00	VIM	0.00	0.72	1.00	0.00	0.00	VIM	0.00	0.15	0.15	0.00	0.00
Snail	-	+	-	+	-	Snail	-	+	-	+	-	Snail	-	+	-	+	-	Snail	-	+	-	+	-
Slug	-	-	+	-	+	Slug	-	-	+	-	+	Slug	-	-	+	-	+	Slug	-	-	+	-	+
TGFB3-Ab	-	-	-	+	+	DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+
OBJ-5	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
B-catenin	0.01	0.99	1.00	0.00	0.25	LEF1	0.01	0.99	1.00	0.00	0.25	LEF1	0.00	0.00	0.00	0.02	1.00	LEF1	0.00	0.00	0.00	0.02	0.65
TCF4	0.87	0.23	0.00	0.77	1.00	E-cad	0.87	0.23	0.00	0.77	1.00	E-cad	1.00	0.04	0.00	0.95	1.00	E-cad	0.93	0.12	0.00	1.00	0.95
Snail	-	+	-	+	-	Snail	-	+	-	+	-	Snail	-	+	-	+	-	Snail	-	+	-	+	-
Slug	-	-	+	-	+	Slug	-	-	+	-	+	Slug	-	-	+	-	+	Slug	-	-	+	-	+
E-cad	-	-	-	+	+	DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+
OBJ-6	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
TGFB3	0.00	0.89	1.00	0.49	0.60	LEF1	0.00	0.14	0.12	0.09	0.09	LEF1	0.00	0.71	1.00	0.31	0.48	LEF1	0.00	0.12	0.16	0.08	0.11
E-cad	0.90	0.00	0.00	1.00	0.71		0.18	0.00	0.00	0.15	0.12		0.52	0.00	0.00	0.92	1.00		0.06	0.00	0.00	0.16	0.17
VIM	0.00	0.97	1.00	0.51	0.66	E-cad	0.00	0.10	0.10	0.05	0.07	E-cad	0.00	1.00	0.94	0.40	0.39	E-cad	0.00	0.15	0.15	0.06	0.06
Snail	-	+	-	+	-	Snail	-	+	-	+	-	Snail	-	+	-	+	-	Snail	-	+	-	+	-
Slug	-	-	+	-	+	Slug	-	-	+	-	+	Slug	-	-	+	-	+	Slug	-	-	+	-	+
E-cad	-	-	-	+	+	DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+	DN LEF1	-	-	-	+	+
OBJ-7	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
Snail	0.00	0.00				Snail	0.00	0.00				Snail	0.00	0.00				Snail	0.00	0.00			
Slug	0.00	0.00				Slug	0.00	0.00				Slug	0.00	0.00				Slug	0.00	0.00			
TGFB3	-	+				Snail	-	+				Snail	-	+				Snail	-	+			
OBJ-8	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
LEF1	0.05	1.00	0.00			LEF1	0.02	0.14	0.00			LEF1	0.02	1.00	0.00			LEF1	0.01	0.25	0.00		
TGFB3	-	-	+			Snail	-	-	+			Snail	-	-	+			Snail	-	-	+		
DN Smad4	-	+	-			Slug	-	+	-			Slug	-	+	-			Slug	-	+	-		
OBJ-9	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
E-cad	0.96	0.00	0.96	1.00		LEF1	0.20	0.00	0.20	0.17		LEF1	0.86	0.00	0.86	1.00		LEF1	0.19	0.00	0.19	0.22	
VIM	0.00	1.00	0.00	0.00		E-cad	0.00	0.11	0.00	0.00		E-cad	0.00	1.00	0.00	0.00		E-cad	0.00	0.24	0.00	0.00	
TGFB3	-	+	-	+		Snail	-	+	-	+		Snail	-	+	-	+		Snail	-	+	-	+	
DN Smad4	-	-	+	-		Slug	-	-	+	-		Slug	-	-	+	-		Slug	-	-	+	-	
DN LEF1	-	-	-	+		DN LEF1	-	-	-	+		DN LEF1	-	-	-	+		DN LEF1	-	-	-	+	
OBJ-10	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
Snail	0.00	1.00	0.00			Slug	0.00	0.19	0.00			Slug	0.00	1.00	0.00			Slug	0.00	0.18	0.00		
Slug	0.00	1.00	0.00			LEF1	0.00	0.20	0.00			LEF1	0.00	1.00	0.00			LEF1	0.00	0.12	0.00		
TGFB2	-	+	+			Snail	-	+	+			Snail	-	+	+			Snail	-	+	+		
DN MEK	-	-	+			Slug	-	-	+			Slug	-	-	+			Slug	-	-	+		
OBJ-11	Mean Trained						S.E. Trained					Mean Predicted					S.E. Predicted				
LEF1	0.05	1.00	0.00			LEF1	0.02	0.15	0.00			LEF1	0.02	1.00	0.00			LEF1	0.01	0.25	0.00		
TGFB2	-	+	+			Snail	-	+	+			Snail	-	+	+			Snail	-	+	+		
DN MEK	-	-	+			Slug	-	-	+			Slug	-	-	+			Slug	-	-	+		

Figure 5.2: Eleven different objective functions used to train our model. Our model was able to effectively predict the simulated experiments 85 percent of the time. The high predictability can be contributed to the leave-one-out cross validation scheme, objective functions with overlapping data, and multi-objective optimization (POETs) algorithm.

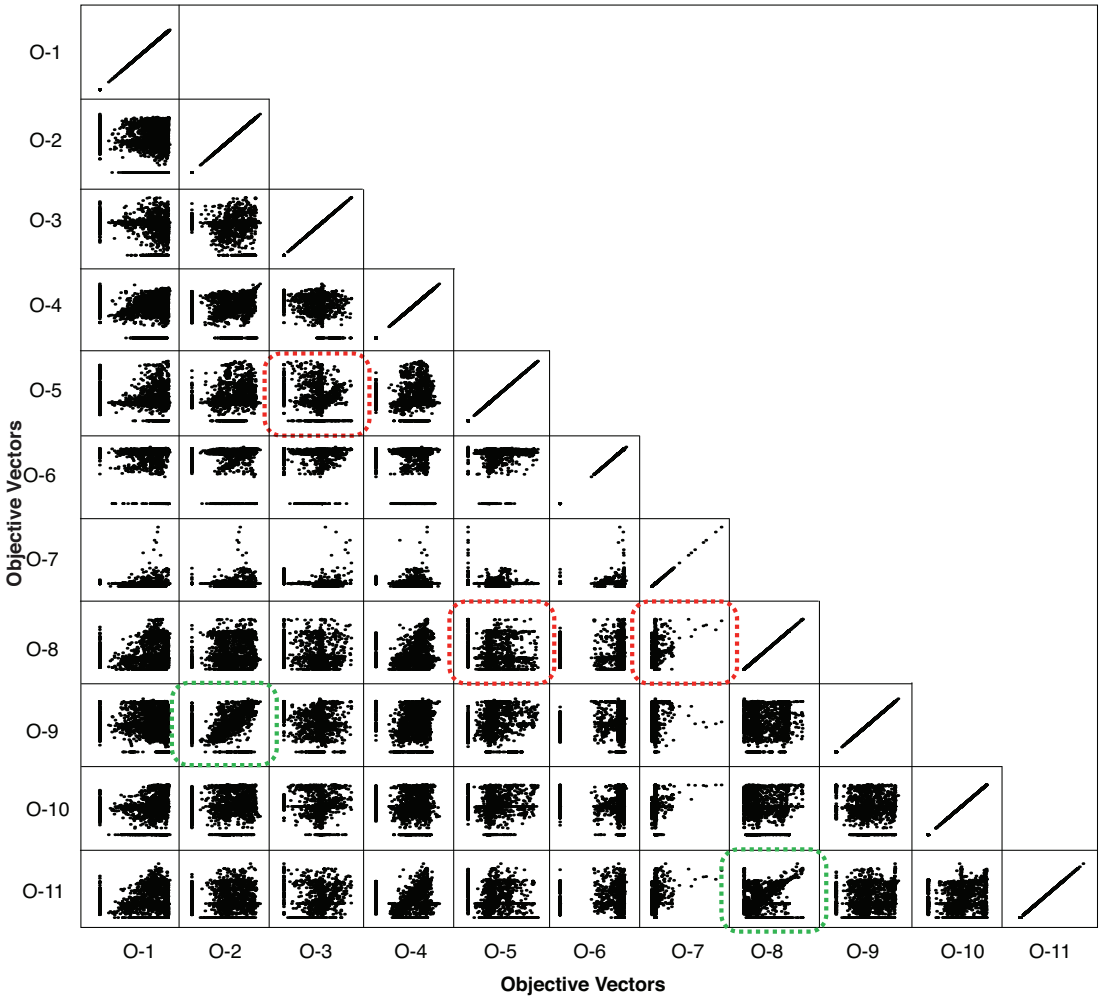


Figure 5.3: Selected objective value plot for the ensemble of EMT models: POETs generated more than 15,000 probable models. We selected models with pareto rank one or less (N = 1091) for further study. POETs was successful in identifying conflicts/relationships between the objective functions which were used for training/predictions. Pareto fronts developed between several objectives, e.g., O3×O5, O7×O8, and O5×O8. These fronts indicated conflicts in the training data and/or an inability of the model to simultaneously capture different aspects of the training data. On the other hand, a linear relationship developed between several objective functions, e.g., O2×O9 and O8×O11. These fronts indicated strong relationships between the training data and was effectively captured by the model .

5.3.3 Signal flow study within EMT model reveals temporal shifts in TGFβ Induced EMT

To investigate temporal shifts for key species dominating the EMT response, we calculated the scaled branch fluxes through the system using steady state vectors. Three modes of operation were simulated to clarify distinct behavioral differences: (a) no dose of TGFβ2 (steady state), (b) TGFβ2 stimulation, and (c) TGFβ2 stimulation while blocking the autocrine response of TGFβ3. Upon TGFβ2 stimulation, MAPK signals through the MEK cascade to activate downstream transcription factors AP1 and SP1. Phosphorylation of these complexes occurs in a sequential order, starting at 1 hr and increases through 48 hrs. When blocking TGFβ3, phosphorylation of these complexes is still maintained (Fig. 5.4A). AP1/SP1 stimulates Snail/Slug repression of the E-cadherin complex to reduce substrate affinity β-catenin and accumulation of free cytosolic β-catenin ensues(Fig. 5.4B). On the contrary, blocking TGFβ3, membrane bound β-catenin increased after 10 hrs, while low levels of free cytosolic β-catenin accumulated between 1-10 hrs (Fig. 5.4C). Accumulation of free β-catenin is critical for complexing with TCF4 to rapidly produce the autocrine response of TGFβ3 between 1-10 hrs. During the autocrine response of TGFβ3, formation and spatial location of the Smad complexes occurs in a time dependent manner. Smad2 phosphorylation occurs within 1 hr, the complex pSmad2/4 forms within 5 hrs, and nuclear localization of the complex dramatically increases after 10 hours(Fig. 5.4D). At steady state, scale branch flux the E-cadherin complex was maintained at a basal epithelial level with no expression of Vimentin. During TGFβ2 stimulation, both MAPK and Smad act synergistically to repress the E-cadherin complex within 1hr. This is followed by increase of Vimentin at 1hr, while expo-

nentially increasing 10 hrs. Both species elevate within 1 hr, however the complex does not significantly form until 10 hrs (Fig. 5.4E). Upon blocking TGFβ3, Snail/Slug downregulates the E-cadherin complex between 5-10 hrs followed by an overwhelming response of TF1 to reverse expression of E-cadherin at 10hr (Fig. 5.4F).

5.3.4 The ensemble of TGFβ Induced EMT models recapitulated the distinct signaling pathways associated with TGFβ isoforms

Our ensemble of parameter sets recapitulates the core signaling pathways associated with TGFβ induced EMT (Fig. 5.5) (Table. 5.2). The literature values used for training purposes were conducted in DLD1 colon carcinoma, MDCKII, and A375 melanoma cells [347]. TGFβ2 signals through the MAPK pathway to upregulate expression of snail and slug and downstream effectors, such as LEF1. This was captured through a (10 a.u.) stimulation of TGFβ2 and inhibition of the MEK complex (Fig. 5.5A-B). To define the role of snail and slug, simulated overexpression studies were conducted(Fig. 5.5C-F). These simulations constrained the effects of snail and slug to indirectly upregulate the autocrine response of TGFβ3 through the β-catenin-TCF4 complex. Overexpression of the E-cadherin complex was able to inhibit the epithelial transformation by sequestering available cytosolic β block EMT. TGFβ3 signals through the Smad pathway to upregulate expression of LEF1 and downstream effectors. This was captured through a (10 a.u.) stimulation of TGFβ3 and inhibitory studies using the DN-Smad4 and DN-LEF1 complex (Fig. 5.5G-I). Expression of both Vimentin and E-cadherin

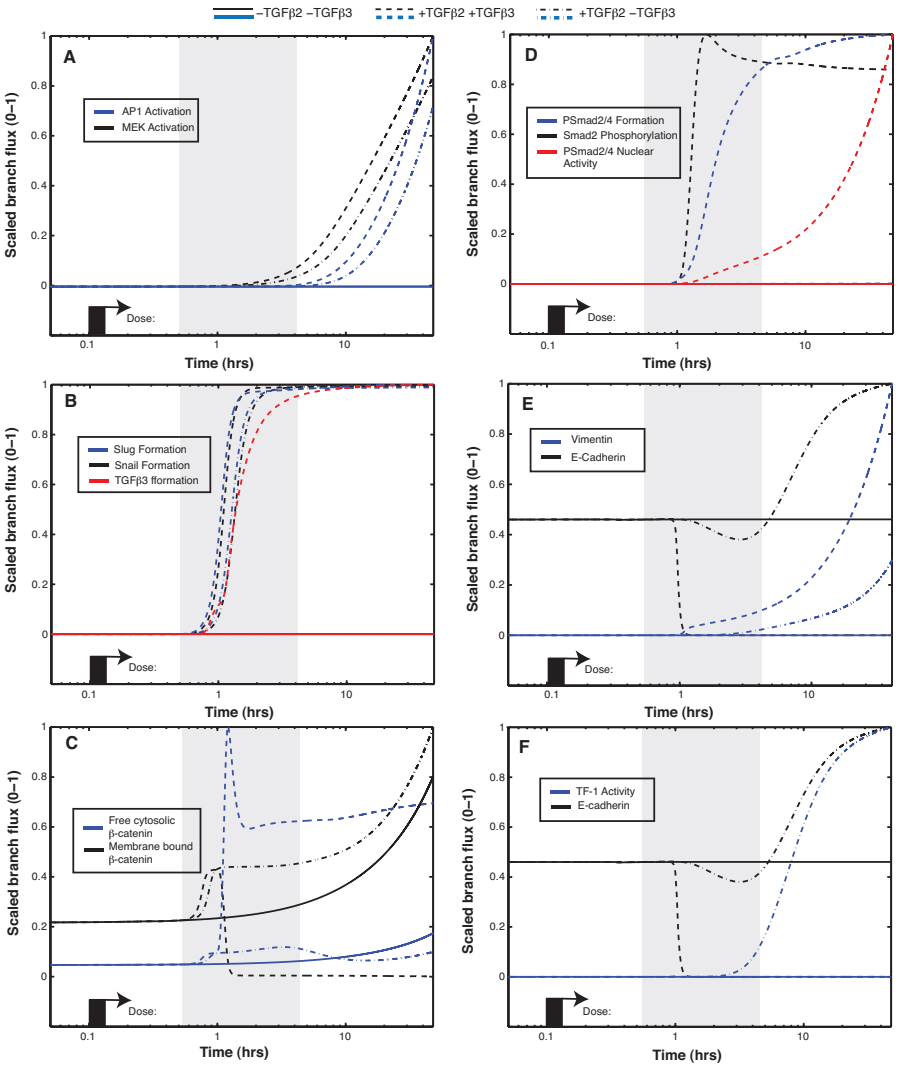


Figure 5.4: Signal flow analysis of key species at steady state, TGFβ2 stimulation, and blocking the TGFβ3 autocrine response. (A) TGFβ2 stimulates the MAPK pathway to activate transcription factor AP1 (5-10hr). (B) The MAPK cascade is directly responsible for rapid expression of snail and slug and downstream TGFβ3 formation (1hr). (C) TGFβ2 reduces membrane affinity for b-catenin, allowing rapid free-cytosolic b-catenin to accumulate (1hr). Blocking TGFβ3 increases membrane bound b-catenin (10hr). (D) TGFβ3 activates the Smad cascade. Nuclear localization of the pSmad2/4 complex (10hr) is dependent upon both the phosphorylation of Smad2 (1hr) and complexing with Smad4 (5hr). (E) TGFβ2 rapidly reduces the E-cadherin complex, while upregulating Vimentin (5-10 hours). Blocking TGFβ3 increases E-cadherin (10hr) and Vimentin is significantly reduced. (F) Expression of the E-cadherin complex (10hr) is positively regulated by the TF-1 transcription factor.

were shown dependent upon these complexes. Although these sets do not encompass the full set of trained data, they provide a core sample of the molecular mechanism driving TGFβ induced EMT. Further sets are included in the supplement. In whole, our model was able to effectively predict the simulated experiments 85 percent of the time. The high predictability can be contributed to the leave-one-out cross validation scheme, objective functions with overlapping data, and multi-objective optimization (POETs) algorithm. The model was also compared against temporal data using MDCKII cells, to measure the effectiveness as a pure prediction [346]. E-cadherin expression reduced exponentially over a 72 hours and was captured within one standard error of our prediction. pSmad2 production increased exponentially over a 72 hours and was also captured within one standard error of our prediction. LEF1 production increased exponentially over a 72 hours, but could only be observed effectively after the 48hr time point(Fig. 5.5J-L).

5.3.5 Population sensitivity analysis revealed globally and temporally important network components

Sensitivity analysis has been widely used to understand fragility within networks [176]. First order sensitivity coefficients were computed for two cases: (a) TGFβ2 stimulation, and (b) TGFβ2 stimulation while knocking down the autocrine response of TGFβ3 using 55 parameter sets selected from the parent ensemble. These 55 parameter sets were chosen based on diversity from the nominal parameter set, its coefficient of variation, and low correlation (Fig. 5.6).

The coefficients were time-averaged (from 0-48 hrs or in pre-selected time

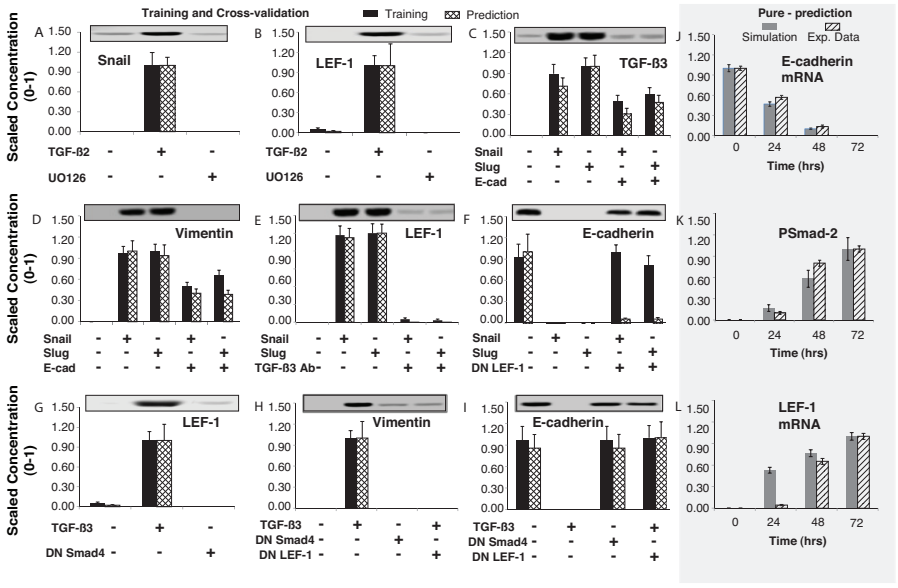


Figure 5.5: Using POETs and cross-validation, our ensemble of parameter sets recapitulates the core signaling pathways associated with TGFβ induced EMT. (A-B) TGFβ2 (10 a.u.) signals through the MAPK pathway to upregulate expression of snail and slug and downstream effectors, such as LEF1. (C-F) Overexpression of snail and slug indirectly upregulated the autocrine response of TGFβ3 through the β-catenin-TCF4 complex. In contrast, overexpression of the E-cadherin complex was able to inhibit the epithelial transformation by sequestering available cytosolic β block EMT. Likewise, DN-LEF1 and TGFβ3-Ab inhibited the epithelial transformation. (G-I) TGFβ3 signals through the Smad pathway to upregulate expression of LEF1 and downstream effectors. This was captured through a (10 a.u.) stimulation of TGFβ3 and inhibitory studies using the DN-Smad4 and DN-LEF1 complex. Expression of both Vimentin and E-cadherin were shown dependent upon these complexes. (J-L) The model was also compared against untrained temporal data, to measure the effectiveness as a pure prediction. E-cadherin expression reduced exponentially over 72 hours, pSMAD2 production increased linearly over a 72 hours and, LEF1 production increased over a 72 hours, but could only be captured after effectively after the 48hr time point.

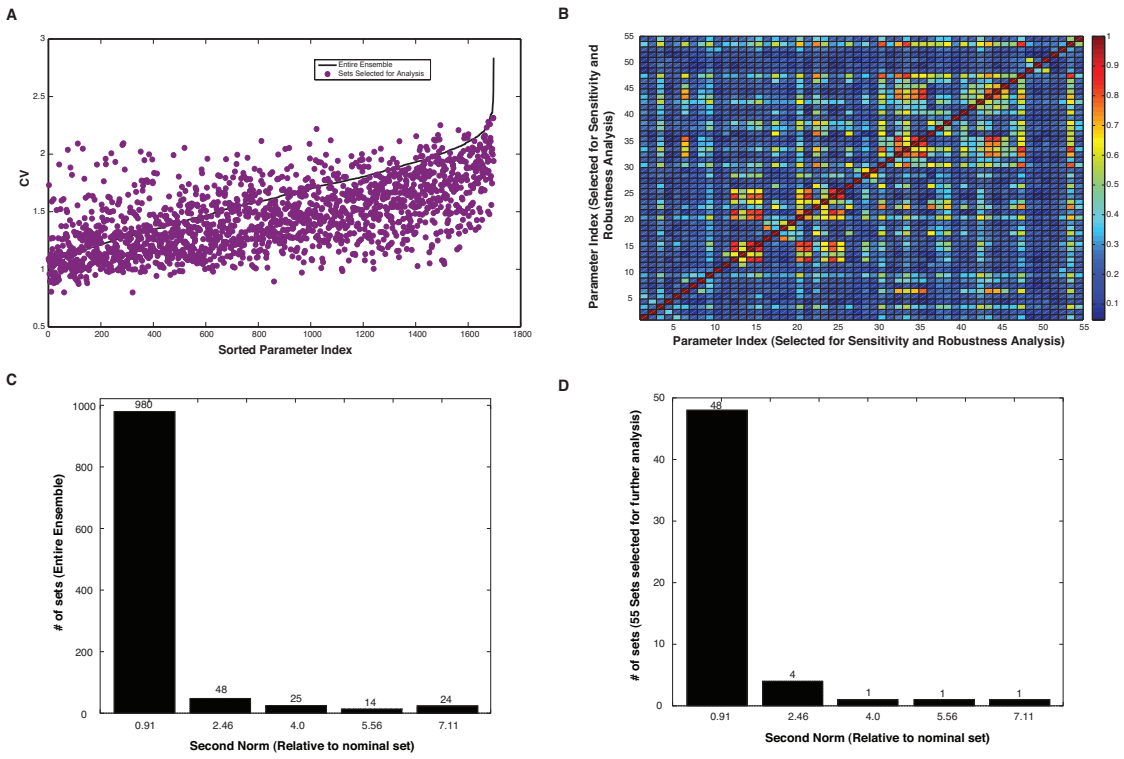


Figure 5.6: Using POETs we generated N=1091. (A) The coefficient of variation (CV) for the parameters in the model ensemble ranged from 0.9 - 2.8 (black dots). 55 parameter sets were chosen for subsequent analysis (purple dots). (B-D) The 55 parameter sets were chosen based on diversity from the nominal parameter set (as seen by the histogram of the second norm in C-D). Approximately % of parameter sets its coefficient of variation, and low correlation.

bands) to form the $\mathcal{N}$ array and decomposed using SVD (materials and methods). The magnitude of the coefficients of the left (right) singular vectors corresponding to largest singular values of $\mathcal{N}$ was used to rank-order the importance of the nodes (edges) in the model. Given the phased behavior of TGFβ isoforms involved in the system, we analyzed the temporal changes in the specie sensitivity within four separate time bins (0-2, 2-5, 5-10, and 10-48 hrs). Deviation from the 45 was used to see qualitative differences in terms of importance (species).

In terms of species for the TGFβ2 perturbed case, mRNA expression levels of E-cadherin were seen to be more important at early (0-2 hrs) compared to later times (10-24 hrs) (Fig. 5.7A). This was expected as snail was rapidly induced early on to repress the E-cadherin gene. However, the actual E-cadherin protein complexes were not affected until later time periods (5-10 hrs). Phosphorylated Smads and inhibited Smads were also seen important later (10-48 hrs) compared to early times (0-2 hrs). In particular pSmad2/4/LEF1 complex and pSmad3/4 were the most important at later times (10-48 hrs). This can partially be attributed to delayed response of nuclear localization of Smad complexes as shown through the flux analysis. Interestingly, AP1 and SP1 complexes were also seen to be important at later times (10-24 hrs). This possibly suggests that sustained activation of the MAPK pathway is needed in conjunction with the Smad pathway to fully induce EMT.

In the case of the blocked response to TGFβ3 autocrine signaling, TF1 mRNA and E-cadherin mRNA levels were seen at early (0-2 hrs) times followed by the formation of the E-cadherin complex (5-10hrs). At later time frames (10-48 hr), activated ERK, SP1, and AP1 were shown to be important (Fig. 5.7B). Interestingly, β-catenin/TCF4 complexes are similar to the TGFβ2 perturbed case. This suggests that snail/slug repression of E-cadherin is important for reducing substrate affinity for β-catenin and accumulation of cytosolic β-catenin. However, full repression of the E-cadherin complex does not occur without the combined effect of the pSmad2/4/LEF1 complex. This highlights the sequential role of both snail/slug and pSmad2/4/LEF1 repression complexes to inhibit E-cadherin formation.

Overall time-averaged sensitivity was conducted to identify the globally im-

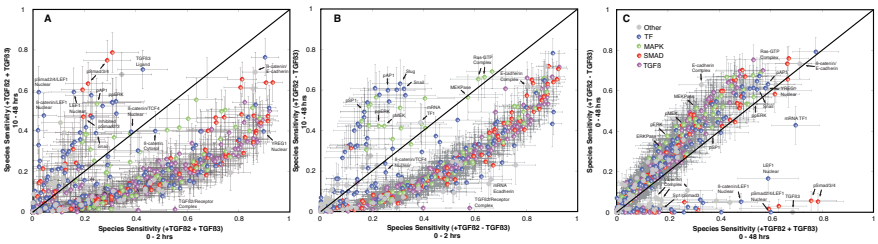


Figure 5.7: First order sensitivity coefficients were computed for the temporal response of TGFβ2 stimulation, and blocking the TGFβ3 autocrine response. (A) During TGFβ2 stimulation pSmad2/4/LEF1 complex and pSmad3/4 were the most important at later times (10-48 hrs) followed by the AP1 and SP1 complexes. (B) When blocking the TGFβ3 autocrine response mRNA expression levels of E-cadherin were seen at early (0-2 hrs) times, however, at later time frames (10-48 hr), activated ERK, SP1, and AP1 were shown to be important. Degradation of pSmad2/3 and expression of TF1mRNAs also transpire. (C) Overall time-averaged sensitivity revealed differential regulation of key complexes. pSmad2/Smad4/LEF1 was most important in the TGFβ2 perturbed case. In both cases, the transcription factor SMAD3 was ranked among the highest followed by membrane bound β-catenin, YREG1, RasGTP, and MAPK phosphatases. This suggests the importance of sustained activation of the both MAPK/Smad pathways, potential negative feedback through the ERK cascade, and pSMAD2/4/LEF1 as a critical complex within the TGFβ2 perturbed case.

portant species in each condition (Fig. 5.7C). In whole, we distinctly see species which are important in +TGFβ2 +TGFβ3 case as compared to +TGFβ2 -TGFβ3. These include components like the pSmad2/4/LEF1 and pSmad3/4 complexes. Less sensitive but also important components include the Smad3/SP1 complex, which has been shown to regulate TGFβR3. In both cases, the cytosolic Smad3 was ranked among the highest followed by membrane bound β-catenin, YREG1, and Ras-GTP. Interestingly, MAPK phosphatases were also important in both cases, suggesting a feedback mechanism to regulate high activated levels of the

ERK cascade.

5.3.6 Robustness analysis revealed hypothetical schemes in regulating phenotypic shifts in TGFβ induced EMT

TGFβ signal transduction networks are notorious for its overwhelming large numbers of interacting species, redundant signaling modules, and crosstalk between modules. These topological features often lead to complex synergistic behaviors as well as robustness [510]. However, networks which are highly optimized for specific tasks may also contain hidden fragility [65]. Here, we generated falsifiable predictions about the fragility or robustness of the EMT network using robustness analysis. Robustness coefficients were used to quantify the effect of structural perturbations (node deletions and over-expressions) on 995 species in the EMT model. Robustness coefficients with values > 1 (< 1) indicated a marker increased (decreased) compared to a base state, while a value of 1 indicated approximately no change following a perturbation.

Gene knockdowns with TGFβ2 perturbation for the 88 elemental species were computed using 55 parameter sets from the ensemble (Fig. 5.8A). These parameters were identical to the sensitivity analysis and selected based on diversity from the nominal parameter set (further details in supplemental materials). Phenotypic behavior of the ensemble was captured using downstream markers Vimentin (mesenchymal) and E-cadherin (epithelial) at 48 hours. This was used to identify the role of individual gene components and understand population average behavior. Overall, the 88 elemental knockouts shifted the population towards an epithelial phenotype. As seen earlier in the sensitivity

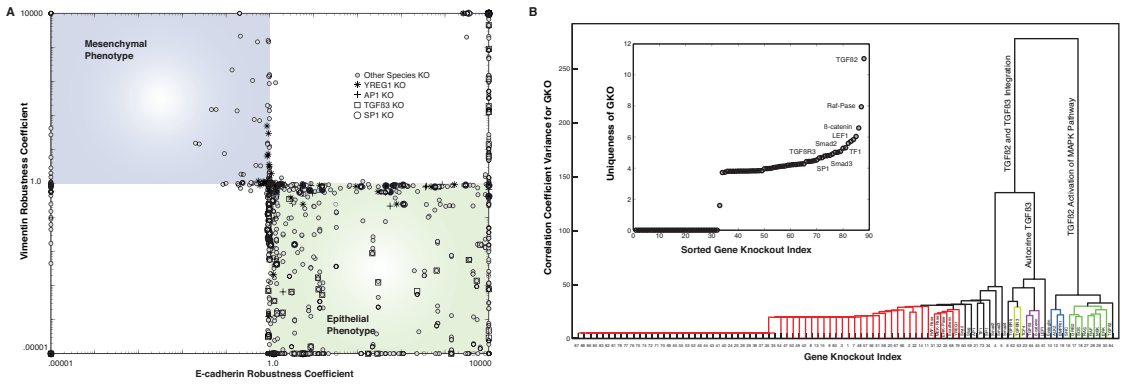


Figure 5.8: Analysis on all 88 elemental genes revealed a key species involved in shifting cellular phenotype. (A) Gene knockdowns with TGFβ2 perturbation for the 88 elemental species were computed using 55 parameter sets from the ensemble. Phenotypic behavior of the population was analyzed using downstream markers Vimentin (mesenchymal) and E-cadherin (epithelial). YREG1, AP1, SP1, and TGFβ3 were just a few of the important species for regulating phenotype. (B) A dendrogram of the single gene knockouts was used to reveal clusters with inherent functional relationships. The most distinct separation was between the receptor components of the MAPK pathway and TGFβ3 signal induction through β-catenin/LEF1.

analysis, key species such as LEF1, TGFβ3 and Smad4 were differentially more important in the presence of TGFβ3 and prevented the shift towards the mesenchymal phenotype. These findings are consistent with the observations in the training data. Other species such as TGFβR3, AP1, SP1, and Smad3/4 were also shown to be important for regulating phenotype. Interestingly, a few knockouts even resulted to increase mesenchymal phenotype, e.g., YREG1.

We further investigated functional classifications within these knockouts. A dendrogram revealed clusters of inherent functional relationships (Fig. 5.8B). The most distinct separation was between the receptor components of the MAPK pathway and TGFβ3 signal induction through β-catenin/LEF1. These

clusters represent the sets of knockouts capable of producing qualitatively similar perturbations in the system. Further clustering revealed similarities between bothTGFβ2 and TGFβ3 knockouts. In particular, Smad2/3, SP1/AP1, TF1, Snail/Slug, and YREG1. The uniqueness of the 88 gene knockouts was generated through orthogonality. We see that 37.5% of the knockouts did not produce any unique result. At the same time, species like RAF-pase (phosphatase for RAF), β-catenin, LEF1, Smad2, and TGFβR3 produced the most unique result. This suggests that knockouts of these components are potentially important for phenotypic shift and provide a basis for further investigation.

5.3.7 Computational analysis identified hypothetical schemes in regulating phenotypic shifts in TGFβ induced EMT

Using signal flow, sensitivity, and robustness analysis, we identified three important areas of regulation. First, AP1/SP1 have near binding sites to the promoter of the snail/slug region and directly regulate expression in MDCK cells. To maintain an epithelial phenotype at steady state, we incorporated a transcription factor-1 (TF1) which positively expresses E-cadherin and is regulated through the MAPK pathway. As shown through our flux analysis, TGFβ2 induces the Ras/MAPK pathway to activate the AP1/SP1 transcription complexes. Rapid expression of Snail/Slug (<1hr) competes with TF1 for regulating the E-cadherin gene. When knocking down TGFβ3, TF1 overwhelms Snail/Slug and dominates E-cadherin expression (>5hr). Thus, the early phase of epithelial transformation is directly related to the regulation (magnitude and binding affinity) of Snail/Slug/TF1 expression. Using robustness analysis, AP1

knockouts revealed distinct subpopulations of parameter sets capable of shifting towards a mesenchymal phenotype. This was shown through negative expression of the E-cadherin complex. Similarly, SP1 knockdowns revealed distinct subpopulations that were capable of positively expressing Vimentin (Fig. 5.9B). Taken together, these results suggested the differential role of AP1/SP1 throughout the ensemble of parameter sets and may prove as an important transformational aspect between cellular phenotype.

Second, over-expression of all 88 elemental species revealed an interesting behavior of the ERK phosphatase. As one would expect, over-expression should reduce the phosphorylated levels of the MAPK cascade and mesenchymal phenotype. It is well known that MPKs provide negative feedback mechanism for MAP kinase activity and experimental over-expression studies have confirmed reductions within the oncogenic-Ras cell lines. However, our model has identified a sub-population of parameter sets that was shown capable of exacerbating the mesenchymal phenotype (Fig. 5.9C). After further investigation, it was revealed that the negative feedback induced by ERK on Smad regulation was the contributing factor. During the over-expression of the ERK phosphatase, a reduction of phosphorylated ERK lead to a larger decline of inhibited Smads compared to MAPK activity. Therefore, within this sub-population of parameter sets, an increase of phosphorylated pSmad2 was shown (Fig. 5.10A). To confirm that the increase of pSmad2 was a function of reduced ERK inhibition and not from another source, we found that the total levels of inhibited pSmad2 through targeted ERK was also less (Fig. 5.10B). As a positive control, we double over-expressed ERK and MEK phosphatase which lead to complete inhibition of the MAPK cascade and mesenchymal phenotype (Fig. 5.10C). In whole, a balance between the ERK cascade and inhibition of Smads is clearly seen. Selected

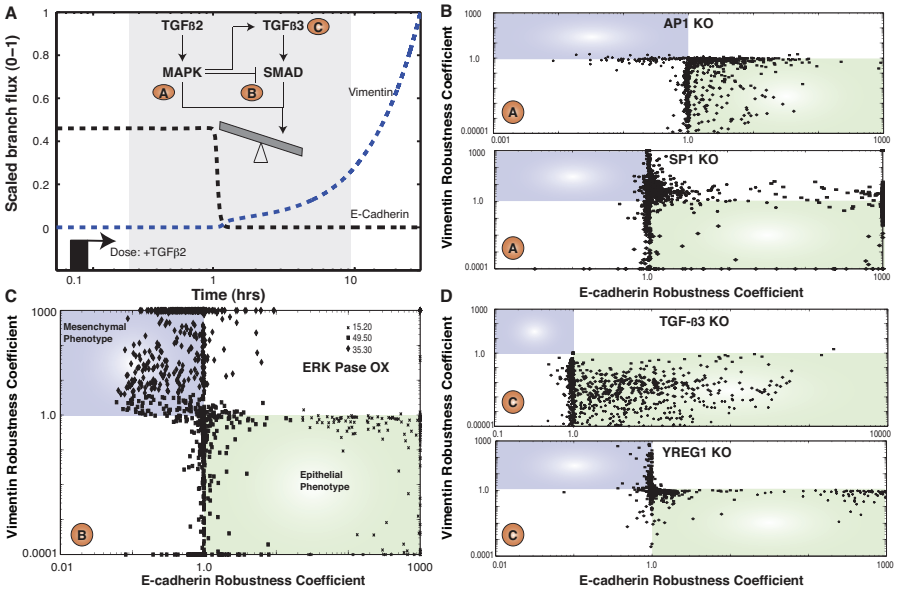


Figure 5.9: Using signal flow, sensitivity, and robustness analysis, we identified three important areas of regulation. (A) The differential role of AP1/SP1 throughout the ensemble of parameter sets may prove as an important transformation aspect between normal and cancerous cells. Both AP1/SP1 knockouts revealed distinct subpopulations of parameter sets capable of shifting towards a mesenchymal phenotype. (B) A balance between the ERK cascade and inhibition of Smads plays an important role in a positive feedback of the TGFβ3 autocrine response. Over-expression of ERK phosphatase lead to a decline of inhibited Smads potentially through a selective dephosphorylation of MAPK/ERK cascade targeting Smad proteins. (C) The availability of LEF1 limits the induction of EMT and takes precedent in forming the Psmad2/LEF1 complex over other complexes. YREG1 knockdowns revealed distinct subpopulations of parameter sets capable of shifting towards a mesenchymal phenotype through expression of Vimentin.

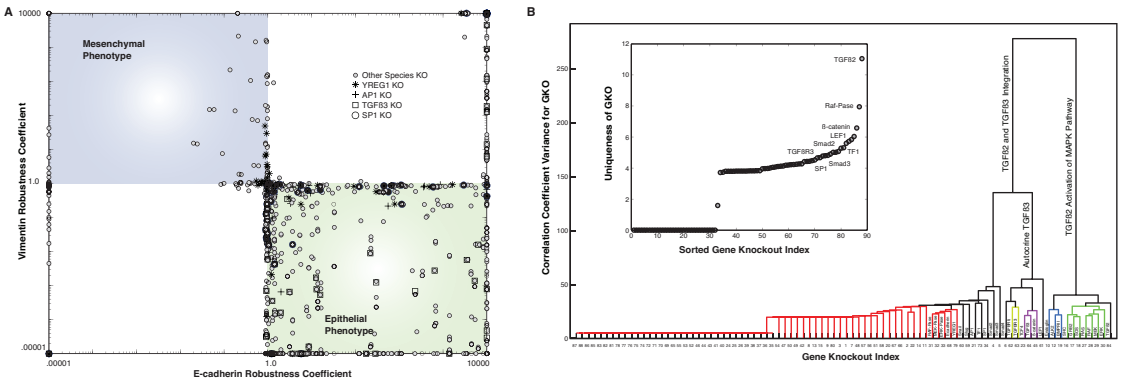


Figure 5.10: Selected phosphorylation through MAPK phosphatases may plan an important role in facilitating a positive feedback mechanism. (A) Sub-population of parameter sets lead to an increase to phosphorylated pSmad2. (B) To confirm that the increase of pSmad2 was a function of reduced ERK inhibition and not from another source, we found that the total levels of inhibited pSmad2 through targeted ERK was less. (C) As a positive control, we double over-expressed ERK and MEK phosphatase which shutdown of the MAPK cascade and thus, mesenchymal shift.

phosphorylation through MAPK phosphatases may play an important role in facilitating a positive feedback mechanism through which the maximum effect of theTGFβ3 autocrine response can occur.

Third, TGFβ 3 knockdowns have revealed a large distribution of parameter sets capable of shifting phenotypic markers. This suggests that production or sensitivity of TGFβ3 may play a role in regulating magnitude of EMT induction. According to our trained data and sensitivity analysis, LEF1 has been shown as an important component downstream of the TGFβ3 autocrine response. Upon TGFβ2 stimulation, DN-LEF1 was capable of inhibiting EMT. Recent work by *Arce et al.* has shown a complex role of (Amino terminal Enhancer of Split) AES and Groucho/TLE on suppressing LEF1 activity. AES opposes LEF1 transcrip-

tional activation while Groucho/TLE binds with LEF1 for a histone deacetylase repression. In addition, β-catenin directly displaces Groucho/TLE repressors from TCF/LEF1 in Wnt-mediated transcription activation [14]. Although we did not implement the direct structure of this complex feedback, we introduced the hypothetical protein YREG1 which directly represses LEF1. YREG1 knock-down and over-expression studies revealed distinct subpopulations of parameter sets capable of shifting towards a mesenchymal phenotype. This was shown through positive expression of the Vimentin complex. Surprisingly, minimal changes occurred within the E-cadherin complex (Fig. 5.9D). Thus, transcriptional complexes such as the the β -catenin/LEF1 complex, which drives Vimentin expression, are much more effected by YREG1. On the contrary, the pSmad2/4/LEF1 complex, which binds to E-cadherin, is minimally effected by YREG1. This suggested that the availability of LEF1 takes precedent in forming the pSmad2/LEF1 complex over other complexes such as β-catenin/LEF1.

5.4 Discussion

Systems insight into complex cellular responses like EMT holds promise in unraveling the molecular components and logic underlying dynamic integration of signals leading to diverse responses. Previous studies in modeling TGFβ, while limited in scope, provided an insightful analysis for the role of receptor dynamics and Smads. In the current study, we formulated and analyzed a family of mechanistic models (1700 kinetic constants and 995 initial conditions) using a detailed molecular framework which described the global induction of EMT through TGFβ isoforms. The models were identified using 45 different experimental data sets taken from literature from DLD1 colon carcinoma, MD-

CKII, and A375 melanoma cell lines [347]. To address parametric uncertainty in ODE based deterministic models, we implemented an ensemble approach using POETs. The ensemble of parameter sets allowed us to identify different operational paradigms within EMT (sub-populations), which provide insightful information into the possible transformational aspects between cellular phenotype. Using model analysis tools like signal flow, sensitivity and robustness analysis, we were able to identify a few key regulatory modes within TGFβ induced EMT.

The ERK-MAPK pathway has been shown to modify TGFβ signaling at multiple levels. One such way is that the activated Ras pathway inhibits the TGFβ-induced nuclear accumulation of Smad2/3, as seen in epithelial cells [633]. ERK kinases have been shown to phosphorylate Smad2 and Smad3 at specific sites in the region linking the MH1 and MH2 domains. These sites are separate from the TGFβ receptor phosphorylation sites which activate Smad nuclear translocation [338]. The effect of interaction between ERKs and Smads is the subject of some controversy, with data suggesting that such an interaction either enhances or inhibits downstream events. A hyperactive Ras pathway has been shown to effectively counteract the antiproliferative activity of TGFβ through attenuation of Smad accumulation in the nucleus [275]. In contrast, Ras signals strongly cooperated with Smads for invasion of human carcinoma cells [385]. The regulation of MPKs (MAPK phosphatases) are highly specific and tightly controlled through dual specificity. Four unique sub-groups have been shown to regulate specific complexes within ERK,JNK, and p38 cascades. It has been suggested that the role of an additional N-terminal domain within the type III subgroup could be to provide a cross-talk between MAPK and other signaling pathways[123]. Interestingly, our model suggests that upon TGFβ stimulation,

high levels of MAPK induces the autocrine response of TGFβ3 and inhibition of pSmads is limited. A possibility is through selective dephosphorylation of the MAPK/ERK complexes targeting pSmads.

The role of mechanistic mathematical modeling in understanding global molecular phenomenons (such as EMT) remains unclear. Structural and parametric uncertainty associated with this form of modeling study is of critical concern. In this study, a molecular framework was implemented from numerous experimental resources to mimic the canonical signaling pathways associated with TGFβ induced EMT. However, in the current version of the model we lack certain modules of signaling within EMT. In particular, the regulation of β-catenin through the GSK3β complex is a widely studied and intricate component of the EMT process. TGFβ 3 signals through PI3K to activate AKT. AKT can then phosphorylate and inactivate the GSK3β complex. GSK3β is an important target of both Snail and β-catenin for degradation through the ubiquitin-proteasome pathway. Thus, it provides a critical negative feedback mechanism which would should be incorporated in further models. Specifically to our model, expression of Snail increases through 72 hours. In contrast, experimental data has shown that activity of Snail peaks at 24 hours. This is possibly through the regulation of GSK3β and should be included in future models. Addressing missing structural components like that of GSK3β could allow us to get a more comprehensive insight into TGFβ signaling. Given the scope of the problem and limitations imposed by structural/parametric uncertainties, we were able to investigate into the mechanistic workings of EMT and present falsifiable hypothesis relevant to manipulating EMT. While this study was limited to EMT, a similar framework could be used to understand and investigate other similar operational modes relevant to biology.

5.5 Materials and Methods

5.5.1 Formulation and solution of the model equations

The EMT model was formulated as a set of coupled ordinary differential equations (ODEs):

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{r}(\mathbf{x}, \mathbf{p}) \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{5.1}$$

The symbol $\mathbf{S}$ denotes the stoichiometric matrix (995×1700). The quantity $\mathbf{x}$ denotes the concentration vector of proteins or protein complexes (995×1). The term $\mathbf{r}(\mathbf{x}, \mathbf{p})$ denotes the vector of reaction rates (1700×1). Each row in $\mathbf{S}$ described a protein or protein-protein complex, while each column described the stoichiometry of network interactions. Thus, the (i, j) element of $\mathbf{S}$, denoted by σ_{ij}, described how protein i was involved in rate j. If $\sigma_{ij} < 0$, then protein i was consumed in r_j. Conversely, if $\sigma_{ij} > 0$, protein i was produced by r_j. Lastly, if $\sigma_{ij} = 0$, there was no protein i in rate j. All of these interactions were obtained from the literature.

We assumed mass-action kinetics for each interaction in the network. The rate expression for interaction q was given by:

$$r_q\left(\mathbf{x}, k_q\right) = k_q \prod_{j \in \{\mathbf{R}_q\}} x_j^{-\sigma_{jq}} \tag{5.2}$$

The set $\left\{\mathbf{R}_q\right\}$ denotes reactants for reaction q while σ_{jq} denotes the stoichiometric coefficient (element of the matrix $\mathbf{S}$) governing species j in reaction q. All reversible interactions were split into two irreversible steps. The mass-action formulation, while expanding the dimension of the EMT model, regularized the mathematical structure; this allowed automatic generation of the

model code using UNIVERSAL and regularized the unknown model parameters (parameters were one of only three types, association, dissociation or catalytic rate constants). UNIVERSAL, an open source Java code generator, generates multiple code types from text and SBML inputs. UNIVERSAL is freely available as a Google Code project (http://code.google.com/p/universal-code-generator/). Thus, although mass-action kinetics increased the number of parameters and species, they reduced the complexity of model analysis. In this study, we considered well-mixed nuclear, cytosolic and extracellular compartments. The model equations were solved using the LSODE routine in OCTAVE (v 3.1.0; www.octave.org) on an Apple workstation (Apple, Cupertino, CA; OS X v10.6.4).

5.5.2 Estimation and cross-validation of a population of models using Pareto Optimal Ensemble Techniques (POETs)

POETs is a multiobjective optimization strategy which integrates several local search strategies e.g., Simulated Annealing (SA) or Pattern Search (PS) with a Pareto-rank-based fitness assignment [504]. POETs has been described previously [504]. The local pattern-search algorithm has been described previously [149, 560]. The parameter ensemble used in the simulation and sensitivity studies was generated from the low-rank parameter sets (Rank ≤ 4) in $\mathbf{K}_i$. We simultaneously calculated training and prediction error during the parameter estimation procedure using leave-one-out cross-validation [261]. The complete set of training data (11 objectives) was subdivided into 11 bins; in each bin 10 data sets were reserved for training while one were reserved for prediction. In the first

bin DS_1 was used for validation while $DS_2 \ldots DS_{11}$ were used for training, etc. Thus, we formulated 11 ensembles from which we evenly selected parameter sets for the *parent* ensemble (supplementary materials). While cross-validation required that we generate additional model populations, we trained and tested against all the data sets.

5.5.3 Sensitivity and robustness analysis of the population of EMT models

Sensitivity coefficients were calculated as shown previously [505, 504, 526] using 55 models selected from the ensemble (supplementary materials). The resulting sensitivity coefficients were scaled and time-averaged (Trapezoid rule) as discussed previously [504]. To estimate the relative fragility or robustness of species and reactions in the network, we decomposed the $N^{(\epsilon)}$ matrix using Singular Value Decomposition (SVD). Coefficients of the left (right) singular vectors corresponding to largest β singular values of $N^{(\epsilon)}$ were rank-ordered to estimate important species (reaction) combinations. Only coefficients with magnitude greater than a threshold ($\delta = 0.1$) were considered. The fraction of the β vectors in which a reaction or species index occurred was used to rank its importance.

5.5.4 Robustness Analysis of the population of EMT models

Robustness coefficients were calculated to understand the robustness of the network as discussed previously [526]. If the robustness coefficient $\alpha > 1$, then the

perturbation *increased* the marker concentration. Conversely, if $\alpha \ll 1$ the perturbation *decreased* the marker concentration. Lastly, if $\alpha \sim 1$ the perturbation did not influence the marker concentration.

5.5.5 Identification of distinguishable knockouts(overexpression) and clustering

Robustness coefficients were used to rank-order knockout(overexpression) experiments in terms of the greatest unique responses and identify species which were linearly distinguishable. The response of the knockout(overexpression) was measured in terms of the robustness coefficients. A dendrogram was derived by considering each of the knockouts(overexpressions) as variables and the average log of robustness coefficient (LRC) for each of the species as observations. The Statistical Toolbox of Matlab (The Mathworks, Natick, MA) was used to generate the distances, linkages and the final dendrogram.

CHAPTER 6

ANALYSIS OF THE EUKARYOTIC HYPOXIC RESPONSE REVEALED HIF-1, AP-1 AND NF-$\kappa\beta$ MEDIATED CROSSTALK BETWEEN IL-8 AND VEGF SIGNALING CASCADES

Authors - Chakrabarti A, Lequieu JP, Kumarga AN, Manzi B, Stojiljkovic M, Chen M, Guros NB, Dromms RA, Tasseff RA, Verbridge SS, Stroock A, Fischbach-Teschl C and Varner JD.

Submitted to - Bio Phys J.

6.1 Abstract

Eukaryotes initiate a complex program in response to low levels of extracellular oxygen. The hypoxic response, which is mediated in part by the action of the HIF1α transcription factor and reactive oxygen species (ROS), involves the integration of several intracellular signaling axes. While the hypoxic response is vital to normal physiology and development, it is also a prevalent feature of angiogenesis. In this study, we developed a family of mechanistic mathematical models of the signaling pathways involved in the eukaryotic hypoxic response. We cataloged the molecular modules mediating different aspects and branches of the hypoxic response and integrated these modules into a single comprehensive model. The integrated hypoxia model (548 molecular species interconnected by 920 interactions), was analyzed to determine critical points of network failure and possible sources of crosstalk and redundancy. Model parameters were estimated by comparing simulations with experimental data. We identified a population of models, consistent with data, using a novel multi-objective

optimization framework combined with cross-validation. EHR model analysis revealed a two-phased response within EHR. Initial activity of NF-$\kappa\beta$ and AP-1 mediated by VEGF and IL-8 signaling cascades lead to autocrine mediated signal amplification ($\sim$ 24-48 hrs) in both VEGF (less significant) and IL-8 (more significant) signals. This phased behavior was validated with experimental studies on MDA-MB231 cells. Structural analysis using extreme pathways identified modes of crosstalk within VEGF and IL-8 signaling e.g., via p38MAPK, MAPK and PKC, and modes of redundancy e.g., divergent sources of transcription factor (NF-$\kappa\beta$, AP-1) activation. On the other hand, sensitivity analysis suggested that regulators of HIF-1α stability e.g., PHD and FIH along with NF-$\kappa\beta$ and AP-1 activation were critical components of the hypoxia program. Computational gene knockdown(overexpression) studies identified network configurations that either amplified or destroyed the hypoxic response and modulated the levels of phenotypic markers. Ultimately, analysis of the model population suggested that interdiction of both NF-$\kappa\beta$ and AP-1 regulatory blocks within EHR was required to disrupt the hypoxic program. These configurations represent experimentally testable hypothesis and potentially new strategies to manipulate the hypoxia program.

6.2 Introduction

Tumors, like healthy tissue, require oxygen and other nutrients to grow [66]. However, for tumors beyond a limiting size, oxygen can not be sufficiently supplied to support continued growth. The eukaryotic hypoxia response (EHR), induced by reduced extracellular oxygen, plays a prominent role in angiogenesis [349, 50]. Angiogenesis is a developmental program important in a vari-

ety of normal physiological contexts [66]. However, in cancer the activation of the angiogenic switch is a critical event in the vascularization of primary and metastatic tumors. The underlying mechanisms controlling angiogenesis are still emerging. Oxygen in the microenvironment is sensed by hypoxia inducible factor 1 α (HIF1α) and the generation of reactive oxygen species (ROS) [300, 355]. HIF1α mediates the initial phase of the angiogenic program by forming a transcriptionally active complex with HIF1β and co-activators such as p300. The stability of the HIF1α subunit is oxygen dependent [429]. In normoxic conditions, hydroxylation at two prolyl residues (P402 and P564) by PHD proteins promotes the association of HIF1α with the Von Hippel-Lindau (VHL) E3 ubiquitin ligase and subsequently leads to degradation. An additional hydroxylation site at N803 near the C-terminus of HIF1α is regulated by the asparaginyl hydroxylase FIH. Hydroxylation at N803 does not influence stability; rather, it blocks the interaction of the HIF1α C-terminal domain with transcriptional co-activators such as p300. Activated HIF1 up-regulates the expression of many factors including VEGF and Interleukin-8 (IL-8) [580]. On the other hand, the second oxygen dependent signaling axis, ROS, promotes nuclear factor κB (NF-$\kappa\beta$) activation [355]. NF-$\kappa\beta$ also regulates both VEGF and IL-8 expression [355, 580]. The exact relationship between ROS and NF-$\kappa\beta$ activation is unclear. However, ROS has been hypothesized to activate a serine kinase which in-turn phopshorylates the N-terminal serine residues (S32/S36) on IKK; potential mechanisms reviewed in [131]. Secreted VEGF and IL-8 signals can then amplify the hypoxic response and induce proliferative signals via both autocrine or paracrine signaling mechanisms through the VEGFR and CXCR receptor families.

Physiochemical models which describe system-level responses can priori-

tize experimental directions, generate testable hypothesis and perhaps identify and validate potential therapeutic targets [59, 16, 259, 256]. However, it is typically impossible to uniquely identify model parameters, even with extensive training data and perfect models [148]. Alternatively, ensemble approaches, which use uncertain model families, have emerged as a promising strategy to deal with model uncertainty in systems biology and other fields like climate prediction [505, 26, 280, 54, 401]. Their central value has been the ability to quantify simulation uncertainty and to experimentally constrain model predictions. For example, Gutenkunst *et al.* showed that predictions were possible using ensembles of signal transduction models despite sometimes only order of magnitude parameter estimates [176]. More recently, Luan *et al.* predicted patient response to therapeutic intervention using an ensemble of human coagulation models [319]. Model ensembles have been used with parameter dependent analysis techniques to robustly estimate important signaling network features. For example, Tasseff *et al.* characterized emergent behavior between androgen and growth factor signaling in prostate cancer using an ensemble of models and sensitivity analysis [526]. Parameter independent techniques such as extreme pathway analysis have also been used to extract qualitatively important features of signaling architectures [406]. Extreme pathways have been used to characterize the structural basis of emergent properties such as crosstalk in signaling architectures [405].

In this study, we developed and analyzed a population of mathematical models of the eukaryotic hypoxic response. The ensemble of EHR models was identified and tested using experimental data from both mouse and human cell lines [601, 267, 336, 343, 264, 312, 241, 170, 454]. Thus, the model described the canonical response of different cell types to oxygen deprivation

and re-oxygenation studies. Model predictions were independently validated by ELISA measurements of secreted VEGFA and IL-8 in MDA-MB231 cells, under hypoxic (1% pO_2) and normoxic (17% pO_2) conditions. Parameter dependent and independent model interrogation techniques identified key aspects and components of EHR signal propagation. Flux-based analysis revealed a two-phased response within EHR. Initial activity of NF-$\kappa\beta$ and AP-1 mediated by VEGF and IL-8 signaling cascades lead to autocrine mediated signal amplification in both VEGF (less significant) and IL-8 (more significant) signals. This phased behavior was validated with experimental studies on MDA-MB231 cells. Structural analysis of EHR using extreme pathway analysis identified modes of crosstalk and redundancy. For example, the highest frequency component involved in crosstalk was RAS-GTP which was involved in the feedback regulation of VEGF and IL-8 cascades via MAPK, while transcription factor activation was highly redundant given the convergent sources of transcription factor controls from both VEGF and IL-8 cascades. On the other hand, sensitivity analysis suggested that regulators of HIF-1α stability e.g., PHD and FIH along with NF-$\kappa\beta$ and AP-1 activation were critical components of the hypoxia program. Computational gene knockdown(overexpression) studies identified network configurations that modulated the levels of phenotypic markers and revealed the basis for emergent properties such as positive and negative synergy. For example, crosstalk between the VEGFA and IL-8 signaling axes was predicted to be mediated by MAPK, PKC and p38MAPK. To successfully manipulate the hypoxic response, it is therefore vital to regulate both NF-$\kappa\beta$ and AP-1 regulatory blocks within EHR. The current EHR model, with inherent structural and parametric uncertainties, identified falsifiable strategies in manipulating the hypoxic response. The EHR model is available in SBML in the

supplemental materials.

6.3 Results

6.3.1 EHR model connectivity.

Model connectivity (human) was assembled from literature and on-line databases; String-8 [233], NetworKIN [305] and TRANSFAC. The EHR model described 548 protein or mRNA species interconnected by 920 interactions (Fig. 6.1). The EHR was modeled using mass action kinetics within an ordinary differential equation (ODE) framework. ODEs and mass-action kinetics are common methods of modeling biological pathways [147, 479, 510, 297, 512, 320, 505, 74, 203].

The model described the integration between extracellular O_2 levels and VEGFA, IL-8 autocrine induced intracellular gene expression, proliferation and death programs. EHR was induced by increased reactive oxygen species (ROS) and increased stability of hypoxia inducible factor-1α (HIF-1α). ROS induction was assumed to be inversely proportional to intracellular oxygen concentration. For HIF-1α, we modeled oxygen-dependent PHD/VHL-induced HIF-1α degradation and FIH-induced inhibition. In addition to HIF-1α, we modeled the expression and regulation of four other transcription factors: AP-1, ATF2, MEF2 and NF-$\kappa\beta$. AP-1 expression was dependent upon both ATF2 and MEF2. Both ATF2 and MEF2 were constitutively expressed and regulated by phosphorylation at T69/T71/T73 for ATF2 and T312/T319 for MEF2 by p38MAPK family members. We modeled a single activating phosphorylation site on AP-

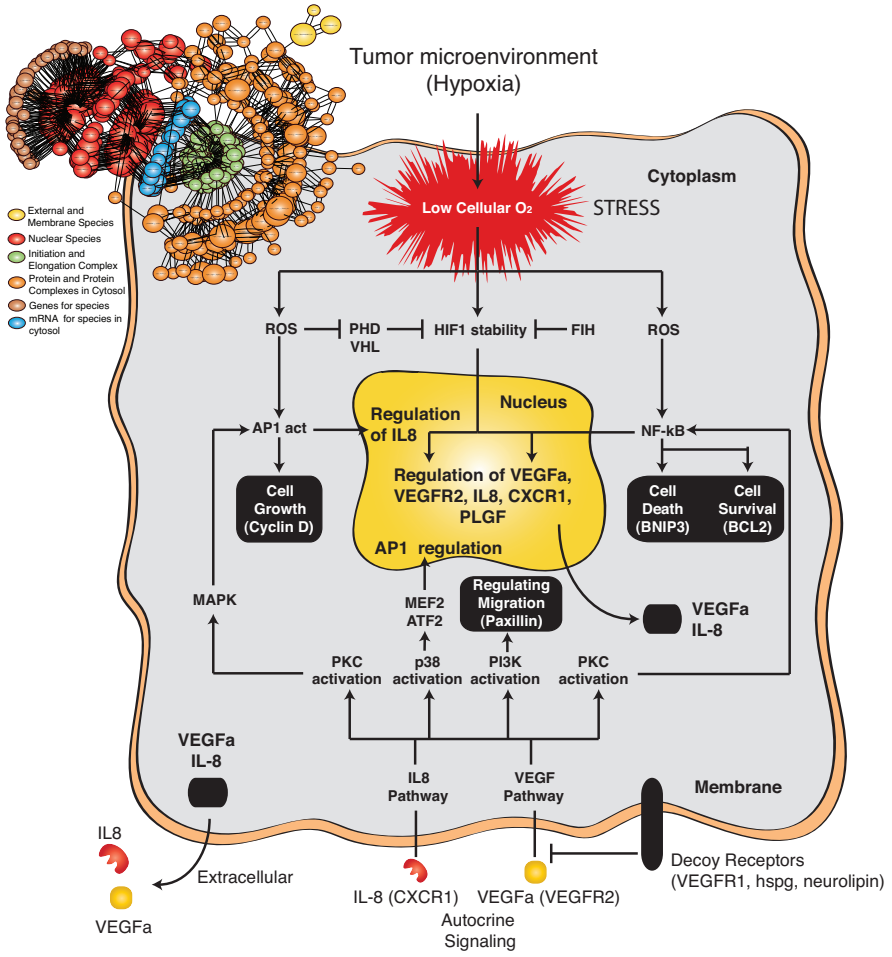


Figure 6.1: The Eukaryotic Hypoxic Response: Low cellular oxygen marks the increased stability of HIF1α and ROS production. Via the role of transcription factors like AP1, NF-$\kappa\beta$ and HIF1, VEGF and IL8 signaling cascades lead to the critical balance mediating tumor (regular cell) growth and angiogenic signaling. Nomenclature: **HIF1** - Hypoxia Inducible Factor 1, **VEGFa** - Vascular Endothelial Growth Factor A, **VEGFR2** - Vascular Endothelial Growth Factor Receptor 2, **IL8** - Interleukin -8, **CXCR1** - IL-8 (or CXCL8) chemokine receptors, **PKC** - Protein Kinase C, **PI3K** - Phosphoinositide 3-kinase, **p38MAPK** - p38 Mitogen-Activated Protein Kinases, **p50_p65** - Nuclear Factor κ -light-chain-enhancer of activated B cells. Species (proteins and protein-protein complexes) and corresponding interactions making up the Eukaryotic Hypoxic Response (EHR) model (Top Left).

1 (S63/S73), which was regulated by ppERK (phosphorylated ERK) which is formed downstream of both IL-8 and VEGFA signaling cascades. NF-$\kappa\beta$ regulation was more complex; NF-$\kappa\beta$ was assumed to be sequestered by a lumped pool of IκB family members. The regulation of this pool occurred by phosphorylation at S32/S36 and Y42 on IκB; S32/S36 was phosphorylated by PKC while Y42 was regulated by an unknown ROS-dependent kinase. Lastly, protein markers downstream of these transcription factors were used to characterize proliferation, survival and death phenotypes. AP-1 dependent Cyclin D expression was used as a marker of proliferation. NF-$\kappa\beta$ dependent Bcl2 expression was used to characterize cell survival while HIF-1α dependent BNIP3 expression was used as a cell death marker. In addition to these markers, the expression of VEGFA, IL-8 and their respective surface receptors VEGFR2 and CXCR1/2 was also modeled. These receptors were connected to transcription factor activation through the activity of the MAPK, p38MAPK and PKC kinases. Complete details of the interactions in the model are enlisted in the supplementary materials.

6.3.2 Estimating a population of EHR models using POETs.

Model parameters were estimated using 20 experimental data sets taken from literature (Table 6.1). In total, 1468 unknown parameters (920 kinetic parameters and 548 initial conditions) were present in the EHR model. Both kinetic parameters and unspecified initial conditions were estimated using the POETs multi-objective optimization algorithm [504]. Additionally, we used cross-validation to independently estimate both the prediction and training error during the parameter estimation (Table 6.1). We used a leave-five-out strategy; five of the twenty objectives were reserved for validation while the remaining 15 were

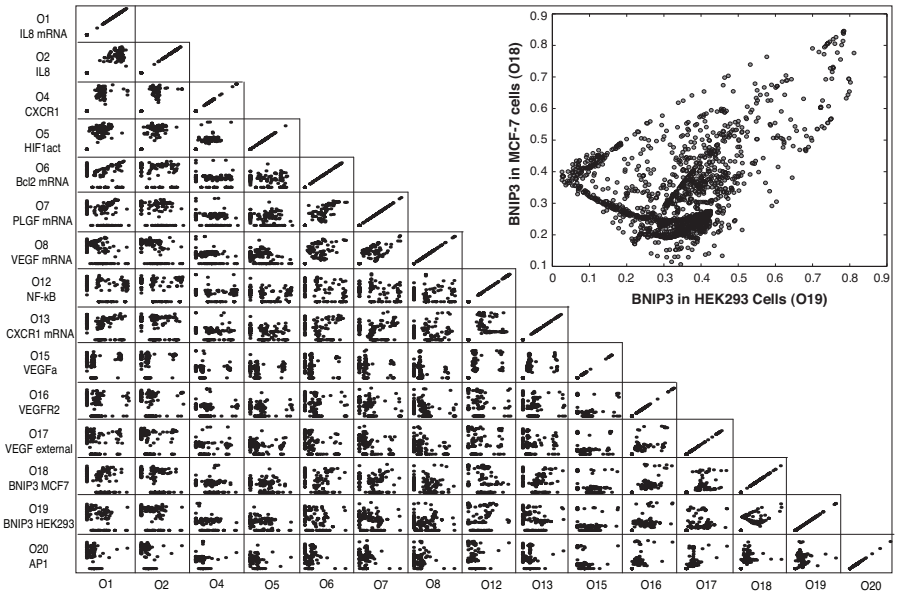


Figure 6.2: Selected objective value plot for the ensemble of EHR models: The grid is a compilation of 2-dimensional plots of the objective values obtained using POETs [504]. Multi-objective optimization scheme allowed us to obtain trade-offs between diverse data sets as seen in O18-O19, O12-O13 and O13-O14. Inset: O18-O19 plot states how we obtain a trade-off between objective functions for BNIP3 in MCF-7 cells and HEK293 cells. This allows us to address qualitative differences in the behavior of BNIP3 in the two cell lines. This figure suggests how POETs was successful in addressing conflicts (due to cell type, experimental protocols etc.) in the data-sets while searching the parameter space.

used for model training. Thus, we estimated four different model families each of which was trained and validated on different experimental data. POETs generated more than 40,000 probable models.

We selected models with Pareto rank one or less (N = 3233) for further study. The coefficient of variation (CV) for the parameters in the model ensemble ranged from 0.8 - 1.8 (supplemental materials). Approximately, 20% of the model parameters were constrained with a CV $\leq$ 1 while 65% of the model parameter sets had a correlation coefficient of less than 80% (supplemental materi-

als). This suggested that we generated a diverse population of EHR models. The most tightly constrained parameters governed oxygen uptake, transcription factor regulation and HIF1α degradation. On the other hand, the least constrained involved the MAPK module and the downstream interactions of BNIP3. Pareto fronts developed between several objectives, e.g., O5×O13, O12×O13, O13×O14 and O18×O19. These fronts indicated conflicts in the training data or an inability of the model to simultaneously capture different aspects of the training data (Fig. 6.2). For example, O18 and O19 were BNIP3 protein measurements taken from MCF-7 and HEK293 cells. A front between these objectives suggested that hypoxia-induced BNIP3 expression varied between these cell-lines. On the other hand, conflicts between CXCR1 expression (O13) and other network components were perhaps rooted in the network topology, e.g., VEGFR and CXCR1/2 mediated competing integration pathways.

The ensemble of EHR models recapitulated diverse training data across multiple cell-lines (Table 6.1). Training data for the EHR model were chosen based upon species connectivity as well as structural analysis of the EHR network. Hypoxic, re-oxygenation and IL-8 induction studies were all used for model training. Approximately, 55% of the models in the ensemble performed better in training versus validation. When compared to the initial best fit model, we lost resolution for ~50% of the objective functions (Table 6.1). The most significant loss was for Bcl2 mRNA (O6) and extracellular IL-8 measurements (O11). On the other hand, we saw improvements for BNIP3 (O19) and NF-$\kappa\beta$ (O12) objective functions.

On average, we predicted VEGFA signaling events better (average prediction error of 0.32) compared to IL-8 (average prediction error of 0.41). A sam-

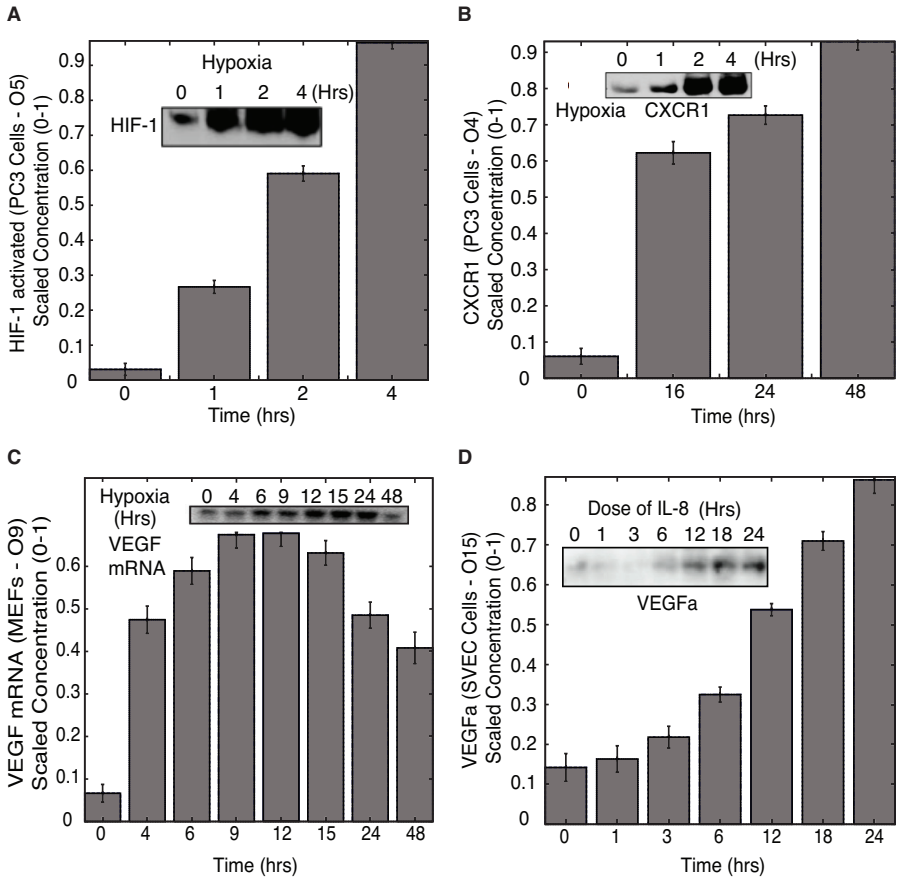


Figure 6.3: Average model performance against data for selected species in the model: Average model performance with one standard error is shown for four key species in the network ((A) HIF-1 data and (B) CXCR1 data used from studies performed by *Maxwell et. al.* on hypoxic PC3 cells [343], (C) VEGFa data used from studies performed by *Romero-Ramirez et. al.* on MEFs [454], (D) VEGFa data used from IL-8 induced studies done by *Martin et. al.* on SVEC cells [336]). Overall the model recapitulated the behavior of species (both upstream and downstream) across diverse data sets [601, 267, 336, 343, 264, 312, 241, 170, 454].

Table 6.1: Objective function list along with species, cell-type, cellular compartment, nominal error (nom), training error, prediction error, random error with a randomly generated parameter set and the corresponding literature reference (ref).

Obj#	species	type	cells	compartment	nom	training	prediction	random	ref
O1	IL-8	mRNA	PC3	Total RNA	0.32	0.45 ± 0.11	0.44 ± 0.13	0.68 ± 0.09	[343]
O2	IL-8	Protein	PC3	Total lysate	0.52	0.50 ± 0.09	0.48 ± 0.08	0.70 ± 0.05	[343]
O3	CXCR1	mRNA	PC3	Total RNA	0.46	0.40 ± 0.11	0.44 ± 0.13	0.73 ± 0.20	[343]
O4	CXCR1	Protein	PC3	Total lysate	0.25	0.30 ± 0.12	0.33 ± 0.11	0.52 ± 0.19	[343]
O5	HIF-1α	Protein	PC3	Total RNA	0.34	0.34 ± 0.13	0.38 ± 0.15	0.66 ± 0.14	[343]
O6	Bcl2	mRNA	PC3	Total lysate	0.09	0.36 ± 0.15	0.37 ± 0.16	0.74 ± 0.06	[343]
O7	PLGF	mRNA	MEF	Total RNA	0.27	0.32 ± 0.12	0.29 ± 0.10	0.75 ± 0.16	[170]
O8	VEGF	mRNA	MEF	Total RNA	0.29	0.28 ± 0.16	0.25 ± 0.15	0.68 ± 0.28	[170]
O9	VEGF	mRNA	MEF	Total RNA	0.42	0.38 ± 0.13	0.31 ± 0.16	0.69 ± 0.09	[454]
O10	HIF1α	Protein	GES	Total lysate	0.25	0.28 ± 0.14	0.25 ± 0.08	0.51 ± 0.23	[312]
O11	IL8 (free)	Protein	HUVEC	Supernatant	0.08	0.33 ± 0.24	0.40 ± 0.27	0.42 ± 0.24	[241]
O12	NF-κB	Protein	Jurkat T	Total lysate	0.72	0.44 ± 0.22	0.42 ± 0.22	0.53 ± 0.18	[264]
O13	CXCR1	mRNA	PC3	Total RNA	0.25	0.39 ± 0.16	0.37 ± 0.17	0.76 ± 0.09	[343]
O14	VEGFa	mRNA	SVEC	Total RNA	0.55	0.43 ± 0.16	0.42 ± 0.19	0.76 ± 0.18	[336]
O15	VEGFa	Protein	SVEC	Total lysate	0.15	0.24 ± 0.23	0.18 ± 0.18	0.58 ± 0.29	[336]
O16	VEGFR2	Protein	SVEC	Total lysate	0.21	0.26 ± 0.19	0.32 ± 0.24	0.57 ± 0.23	[336]
O17	VEGFa	Protein	SVEC	Supernatant	0.19	0.28 ± 0.24	0.35 ± 0.27	0.56 ± 0.30	[336]
O18	BNIP3	Protein	MCF-7	Total lysate	0.34	0.34 ± 0.12	0.33 ± 0.12	0.64 ± 0.12	[267]
O19	BNIP3	Protein	HEK293	Total lysate	0.52	0.26 ± 0.16	0.24 ± 0.13	0.46 ± 0.26	[267]
O20	AP1	Protein	HT29	Total extracts	0.08	0.14 ± 0.11	0.16 ± 0.10	0.81 ± 0.22	[601]

ple of the model performance for key species is shown in Fig. 6.3. The model population captured the enhanced stability of HIF1α following the shift from normoxia to hypoxia ($\leq$ 2% pO2) in PC3 cells (Fig. 6.3A). The enhanced stability occurred over a short timescale (approximately four hours). After translocation, nuclear HIF1α was then free to complex with HIF1β and p300 to drive the initial hypoxia-induced expression program. This second phase occurs on a longer timescale, 24 48 hrs after the shift to hypoxic environment ($\leq$ 2% pO2). The model also captured the long timescale behavior e.g., CXCR1 (Fig. 6.3B) and VEGFA (Fig. 6.3C) expression. VEGFA and CXCR1 expression was regulated by both HIF1α and NF-$\kappa\beta$. Thus, both VEGFA and CXCR1/2 expression is partially driven by positive feedback by VEGF and IL-8 mediated autocrine

signaling. Another example of crosstalk was the connection between exogenous IL-8 and VEGFA expression (Fig. 6.3D). The model recapitulated IL-8 (100 ng/ml) induced VEGFA expression in SVEC cells in a normoxic environment. It suggested this integration occurred following the phosphorylation of IκB at S32/S36 in a PKC-dependent manner, leading to NF-$\kappa\beta$ activation.

We analyzed the dynamics within the EHR network by looking at the flux distributions over time, upon induction of hypoxia ($\leq$ 2% pO2) and upon recovery from hypoxia by inducing normoxic conditions ($\geq$ 17% pO2) (Fig. 6.4A). At hypoxic conditions, we found that there was a two-phased response demarcating the onset of a later autocrine mediated signal amplification. We found that VEGF signaling branch was the first to respond, mediated by the activity of HIF-1 and ROS mediated NF-$\kappa\beta$ activation. Around 24 hours into hypoxia, we saw a second increment in VEGF and IL-8 signaling. This late phase behavior was a result of the positive feedback mediated by the activation of AP-1 and NF-$\kappa\beta$ downstream of both VEGF and IL-8 signaling cascades. Simulated KO studies of AP-1 and NF-$\kappa\beta$ supports their role in signal amplification (supplementary materials Fig. 2).

Comparatively, the dominant increase in the IL-8 signaling branch was more in the second phase of the response (consistent with experimental data as discussed later). However upon induction of normoxic conditions, we did not see a difference in terms of recovery time between VEGF and IL-8 signaling cascades. HIF-1 and NF-$\kappa\beta$ activity dropped down within a few hours of induction of normoxic conditions. On the contrary VEGF and IL-8 took around 10-15 hrs to recover. Steady state flux comparisons (P1 and P2 in Fig. 6.4A-B) revealed that in the ~1% O_2 conditions, flux through the MAPK module is higher compared

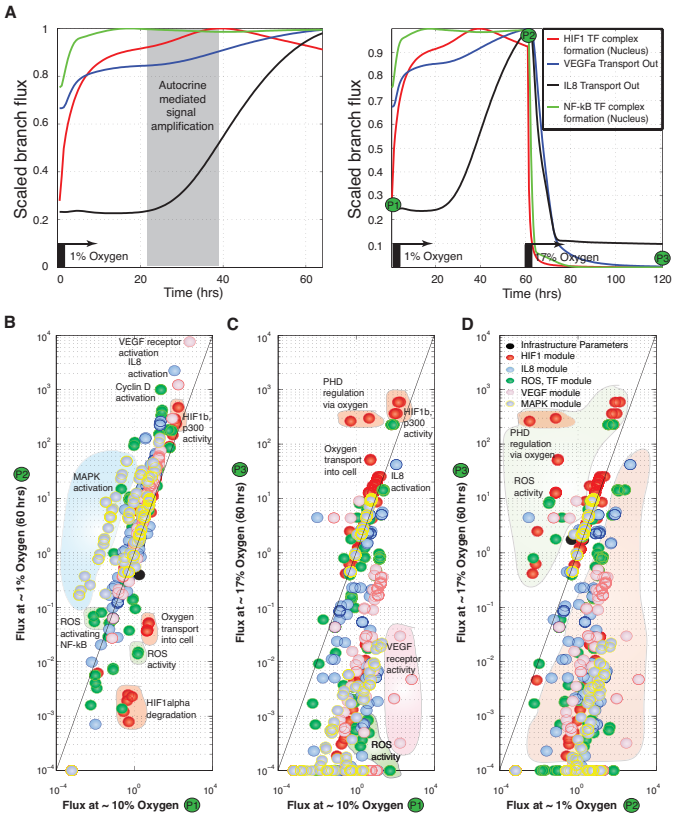


Figure 6.4: Flux analysis of EHR: (A) Plot of the scaled flux (0-1) of HIF-1/NF-$\kappa\beta$ transcription factors and secreted VEGFA/IL8 over time. At hypoxic conditions, there was a two-phased response demarcating the onset of a later autocrine mediated signal amplification. VEGF signaling branch was the first to respond, mediated by the activity of HIF-1 and ROS mediated NF-$\kappa\beta$ activation. Around 24 hours into hypoxia, we saw a second increment in VEGF and IL-8 signaling. However upon induction of normoxic (~17% O_2) conditions, we did not see a difference in terms of time of recovery between VEGF and IL-8 signaling cascades. HIF-1 and NF-$\kappa\beta$ activity dropped down within a few hours of induction of normoxic conditions. On the contrary VEGF and IL-8 took around 10-15 hrs to recover. (B-D) Fluxes at P1 (~10% O_2) when compared to P2 (~1% O_2 after 60 hrs) show increased VEGF and IL8 signaling leading to enhanced MAPK activity along with increased NF-$\kappa\beta$ activity both via ROS and downstream of receptor signaling cascades.

to ~10% O_2 conditions (Fig. 6.4B). This was owing to the convergent MAPK activation downstream of both VEGF and IL-8 signaling. Relatively, there was a hundred fold increase in Cyclin-D regulation in hypoxia (~1% O_2), mediated by activation of AP-1 signaling. Steady state flux comparisons (P2 and P3 in Fig. 6.4A-D) revealed that there was a switch like behavior in the EHR. This switch characteristic was exhibited via the oxygen mediated increased stability of HIF-1α via PHD and ROS activity. For example, components above the 45 in Fig. 6.4D, indicate the dominant mode of operation in normoxic (~17% O_2) conditions. These include PHD and ROS regulation via oxygen, which induces low amount of VEGF and IL-8 signaling. On the contrary, components below the 45 in Fig. 6.4D, indicate the dominant mode of operation in hypoxic (~1% O_2) conditions. These include the VEGF and IL-8 mediated signal integration. Taken together, the population of EHR models recapitulated both short- and long timescale behavior following hypoxia or exposure to exogenous IL-8 in a variety of cell types. It also captured the integration between multiple pathways and generated specific and testable hypothesis about the role of network components in signal propagation.

6.3.3 Model Prediction against 1% and 17% O_2 experimental data in MDA-MB231 cells.

Micro-fabricated 3-D tumor models offer the potential to recreate cell-microenvironment interactions that recapitulate human tumor behavior [563]. Recently, they have been used to investigate the individual contributions of O_2 concentration, culture dimensionality, cellECM interactions, and the coupling

of these effects on tumor angiogenesis [563]. In this study, we measured VEGF and IL-8 secretions of MDA-MB231 cells in alginate hydrogels under hypoxic (1% O_2) and normoxic (17% O_2) conditions at days 1, 3 and 5 (details in materials and methods) and compared it to simulation predictions (Fig. 6.5).

At hypoxic (1% O_2) conditions, we found that secreted VEGF levels almost reach steady state values by day 1 as compared to day 3 and 5 (Fig. 6.5A), which was consistent with model predictions. This behavior indicated that the major fold change in the VEGF signaling occurred in the first phase of the hypoxic response as predicted by the model simulations. However at normoxic (17% O_2) conditions, we saw a decrease in VEGF levels at day 3 and 5 as compared to day 1. We found that the delay in the response of the system to normoxic conditions was of the order of a day, after which the levels dropped down to $\sim$ 50% of the actual values. This behavior could be a result of the reduced levels of HIF-1 and ROS mediated VEGF and IL-8 signal activation. In case of IL-8, at hypoxic conditions we found experimentally that secreted IL-8 levels rose slowly and reached steady state values around day 3 to day 5 (Fig. 6.5B). However model predictions showed a faster rate for IL-8 production. This discrepancy could be due to the lack of important structural information. For example, recently it was shown that HIF-1 induction attenuated Nrf-2 dependent IL-8 expression in HUVECs [315]. Lack of this negative control on IL-8 expression via HIF-1, could be one of the many reasons for the high predicted rate of IL-8 secretion in the model. Interestingly, it was observed both experimentally and predicted by the model that under normoxic conditions secreted IL-8 levels increased with time. One potential explanation is that, normoxic conditions leads to even lower amount of basal HIF-1 activation and subsequently more Nrf-2 mediated IL-8 expression. However the model lacked this connectivity, yet it recapitulated this

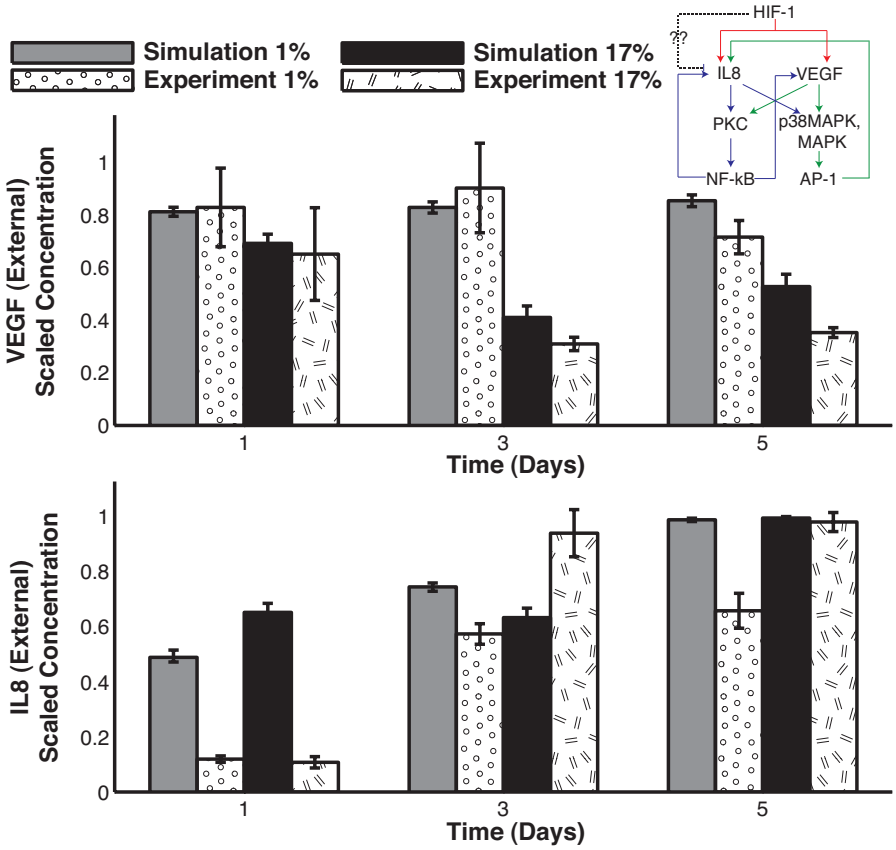


Figure 6.5: Comparison of experimental data (ELISA measurements) versus simulation predictions for secreted VEGF and IL-8 in 1% and 17% O_2 samples in MDA-MB231 cells: At hypoxic (1% O_2) conditions, secreted VEGF levels almost reach steady state values by day 1 as compared to day 3 and 5, indicating that the major fold change in the VEGF signaling occurs in the first phase of the hypoxic response consistent with the prediction by the model simulations. However at normoxic (~17% O_2) conditions, we see a decrease in VEGF levels at day 3 and 5 as compared to day 1, which is a result of reduced levels of HIF-1 and ROS mediated VEGF activation. In case of IL-8, at hypoxic conditions we found experimentally that secreted IL-8 levels rose slowly and reached steady state values around day 3 to day 5. However model predictions showed a faster rate for IL-8 production. Interestingly, it was observed both experimentally and predicted by the model that under normoxic conditions secreted IL-8 levels increased with time.

behavior. This was a result of lower amount of CXCR1/2 production leading to higher amount of free extracellular IL-8 levels. Taken together, the population of EHR models with inherent structural and parametric uncertainty provided valid predictions for hypoxic and normoxic experiments.

6.3.4 Parameter independent topological analysis identified the structural basis of EHR crosstalk and redundancy.

Signal transduction architectures frequently contain redundancy, feedback and crosstalk. These topological features ensure signal propagation is adaptable, efficient and robust. However, they also make reprogramming signal flow challenging. To understand the structural basis of crosstalk and redundancy in EHR, we calculated the extreme pathways (EPs) of the EHR network. Extreme pathways (EPs) are systemic pathway vectors which describe steady-state flux distributions through molecular networks. While EPs tell us nothing about network dynamics, the convex combination of these vectors provides a means of understanding possible steady-state phenotypes. Thus, extreme pathways and elementary modes [482] have been used previously to understand the qualitative features of metabolic and signaling networks [511, 406, 405].

To calculate EPs, we used a flux based cut-off (Flux (AU/gdw_hr) > 0.1) to select which interactions were participating at steady state for 10%, 1% and 17% O_2 conditions (points P1, P2 and P3 in Fig. 6.4A). The EPs (2515 for 10% O_2, 2332 for 1% O_2 and 2556 for 17% O_2) of the EHR network were calculated using ExPA [31] and the Octave programming environment (www.octave.org). The structural importance of any interaction was measured on the basis of its

participation in the EPs at any state. In case of 10% O_2, HIF-1β regulation had the most participation number in EPs (~300 EPs) (Fig. 6.6A-B).

VEGFR1 signaling was seen to be more significant in 10% O_2 conditions as compared to 1% O_2 conditions, giving importance to the decoy signaling role of VEGFR1. While in case of 1% O_2, MAPK regulation mediated by MEK-Pase (phosphatase for MEK) and VEGFR2-mediated signaling was seen to be the most important in terms of participation in EPs (~300 EPs). In case of 17% O_2, HIF-1α regulation via O_2 was the set of reactions with the most participation (~170 EPs) (Fig. 6.6A-C). Overall, VEGF signaling elements were structurally most important in both 10% and 1% O_2 conditions. This corroborated with the dominant modes of operation (as seen by flux based studies) at different O_2 levels.

Singular value decomposition (SVD) of the stoichiometric [121] and EP matrices [428] both generated a singular value spectrum where the largest mode had a fractional weight of only 2%. Compared to other biological networks e.g., genome scale *E.coli* and *S.cerevisiae* metabolic networks [448, 137], or the human coagulation network [319], the EHR network was less organized around a single dominant mode. However, the EHR network did contain significant crosstalk and redundancy. Crosstalk and redundancy amongst the EHR network EPs (for 1%, 10% and 17% O_2 conditions) was calculated according to the scheme of Papin and Palsson [406] (Fig. 6.6D). For both 1% and 10% O_2 conditions, we found that ~60% of the EPs had a disjoint output sets and a similar/same input sets, suggesting a divergent nature of the network. For example, oxygen sensing lead to HIF-1 and ROS mediated activation of VEGF and IL-8 signals, subsequently leading to multiple outputs like active PI3K, active MAPK components, active

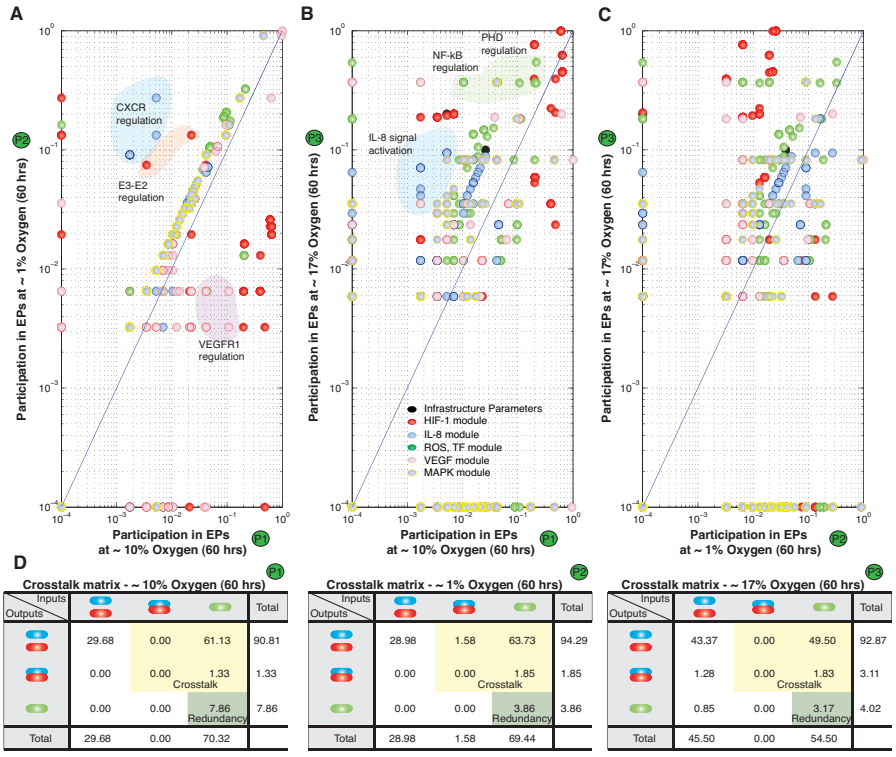


Figure 6.6: Structural analysis of the EHR model: EHR model was analyzed by the Extreme pathway analysis (Details in Materials and methods) to identify the critical structural components. (A-C) Extreme pathway analysis revealed the VEGF module as structurally most important followed by the transcription factor (AP1 and NF-$\kappa\beta$) module. (D) Analysis of the EP's of the EHR network showed presence of significant crosstalk and redundancy. For both 1% and 10% O_2 conditions, we found that ~60% of the EPs had a disjoint output sets and a similar/same input sets, suggesting a divergent nature of the network. For example, oxygen sensing lead to HIF-1 and ROS mediated activation of VEGF and IL-8 signals, subsequently leading to multiple outputs like active PI3K, active MAPK components, active transcription factors e.g. NF-$\kappa\beta$, AP-1. We found that ~5% of the EPs exhibit redundancy. This exhibited redundant routes within the EHR, for example: VEGF could lead to AP-1 activation either directly via MAPK or via p38MAPK. The crosstalk and redundancy observed in the EPs could be the structural basis of complex behaviors such as synergy. On the other hand, the redundant EPs identified the structural basis of EHR robustness, i.e., which interactions could be removed while simultaneously maintaining network function.

transcription factors e.g. NF-$\kappa\beta$, AP-1. We found that ~5% of the EPs exhibit redundancy. This exhibited redundant routes within the EHR, for example: VEGF could lead to AP-1 activation either directly via MAPK or via p38MAPK. Under 17% O_2 conditions, we found that ~50% of the EPs exhibited a divergent nature. At any mode of operation >50% of the EPs encoded network crosstalk, i.e., the same or similar inputs produced variable outputs (Fig. 6.6D). Taken together, the crosstalk and redundancy observed in the EPs could be the structural basis of complex behaviors such as synergy. On the other hand, the redundant EPs identified the structural basis of EHR robustness, i.e., which interactions could be removed while simultaneously maintaining network function. The EP matrix is available in the supplemental materials.

6.3.5 Population sensitivity analysis revealed globally and temporally important network components.

Sensitivity analysis has been used previously in several studies to estimate critical features of networks important to human health, see e.g., [74]. First order sensitivity coefficients were computed for both 10% O_2 and 1% O_2 conditions using parameter sets selected based on CV (supplementary materials). These coefficients were then time-averaged (from 0-30 hrs or in 6 hr time steps) to form the $\mathcal{N}$ array and decomposed using SVD (materials and methods). The magnitude of the coefficients of the left (right) singular vectors corresponding to largest β singular values of $\mathcal{N}$ were used to rank-order the importance of the nodes (edges) in the model (Fig. 6.7 and supplementary materials).

Given the phased behavior involved in the system, we analyzed the tempo-

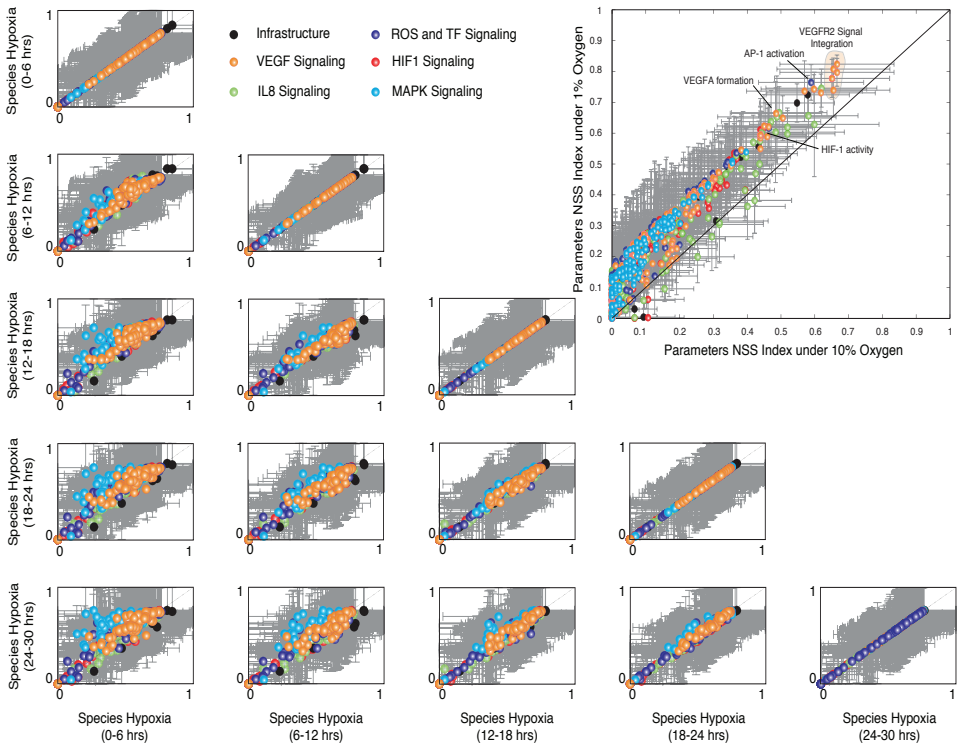


Figure 6.7: Ranked NSS index for species and parameters in the EHR models: Temporal changes in the rank of the species is plotted for 5 time bands (0-6, 6-12, 12-18, 18-24 and 24-30 hrs) in hypoxic (~1% O_2) case. In terms of species in the case of 1% O_2, MAPK components were seen to be more important at later (24-30 hrs) time band as compared to earlier (0-6 hrs) times of EHR. MAPK is a critical regulator of AP-1 activation, which in turn plays an important role at later times (> 24 hrs) leading to autocrine mediated IL-8 signal amplification. AP-1 and NF-$\kappa\beta$ regulatory module was the second group of species which was relatively more important at later time points. Inset - Ranks of the parameters are plotted with one standard deviation for ~1% against ~10% O_2 case. In terms of parameters, infrastructure parameters e.g. nuclear transport, RNA polymerase or Ribosome binding were globally critical components. Interactions governing the regulation of VEGFR2 and VEGFA were the most fragile in both 1% and 10% O_2 conditions. Regulation of the expression and activation of transcription factors like AP-1, HIF-1 and NF-$\kappa\beta$ was in the top 25 fragile species and parameters. In terms of a modular comparison, the transcription factor module was the most fragile, closely followed by the HIF1, VEGF and IL8 modules. Comparatively, the MAPK module was relatively robust.

ral changes in the sensitivities of the species(parameters) in five time bands (0-6, 6-12, 12-18, 18-24 and 24-30 hrs) (Fig. 6.7 and supplementary materials). Deviation from the 45 was used to see qualitative differences in terms of importance in species(parameters). In terms of species in the case of 1% O_2, MAPK components were seen to be more important at later (24-30 hrs) time band as compared to earlier (0-6 hrs) times of EHR (Fig. 6.7). MAPK is a critical regulator of AP-1 activation, which in turn plays an important role at later times (> 24 hrs) leading to autocrine mediated IL-8 signal amplification. AP-1 and NF-$\kappa\beta$ regulatory module was the second group of species which was relatively more important at later time points. This could be attributed to the regulatory role played by AP-1 and NF-$\kappa\beta$ in influencing the positive feedback downstream of both VEGF and IL-8 signaling cascades. Interestingly, both VEGF and IL-8 signaling modules are relatively equally fragile(robust) throughout EHR. In terms of parameters, infrastructure parameters e.g. nuclear transport, RNA polymerase or Ribosome binding were globally critical components. Interactions governing the regulation of VEGFR2 and VEGFA were the most fragile in both 1% and 10% O_2 conditions. This supports the structural importance of VEGF signaling seen by the extreme pathway analysis and the dominant role played by these interactions as evidenced by the Flux analysis. Regulation of the expression and activation of transcription factors like AP-1, HIF-1 and NF-$\kappa\beta$ was in the top 10 fragile species and parameters (Fig. 6.7 Inset and supplementary materials). This result can be attributed to critical nature of AP-1 and NF-$\kappa\beta$ mediated crosstalk in the EHR model via regulation of the VEGF and IL-8 signaling cascades. So the feedback and signal amplification mediated by AP-1 and NF-$\kappa\beta$ was the most critical aspect of the EHR. Other key components identified by sensitivity analysis were related to HIF-1α, IL8 expression, regulation of PI3K and MAPK. Sensitivity

analysis also revealed the fragility of the downstream regulation of Cyclin D and Bcl2. A modular comparison of importance of the interactions revealed the transcription factor (AP-1 and NF-$\kappa\beta$ regulation) module as the most fragile while the MAPK module was relatively robust (Fig. 6.7 Inset). Taken together, sensitivity analysis revealed that while overall MAPK module was relatively robust in the system, at later (>24 hrs) into hypoxia MAPK components are relatively fragile. Overall VEGF signal integration, AP-1/NF-$\kappa\beta$/HIF-1α regulation was seen as the most fragile aspect of EHR. Comparatively IL-8 signaling module was relatively robust.

6.3.6 Robustness analysis presented NF-$\kappa\beta$ and AP-1 as the key regulatory nodes in manipulating the EHR.

Signal transduction architectures frequently contain redundancy, feedback and crosstalk. For example, nearly 70% of the EPs of the EHR network involved crosstalk or redundancy. These topological features often lead to complex synergistic behaviors as well as robustness. However, networks which are highly optimized for specific tasks, may also contain hidden fragility [65]. Here, we generated falsifiable predictions about the fragility or robustness of the EHR network using robustness analysis. Robustness coefficients were used to quantify the effect of structural perturbations (edge/node deletions and overexpression) on network markers. Coefficients with values > 1 (< 1) indicated a marker increased (decreased) compared to a base state, while a value of ~ 1 indicated approximately no change following a perturbation. For each type of perturbation, we calculated the robustness coefficients using the nominal parameter set.

We calculated the direction of O_2-induced concentration shifts for 548 markers following single parameter knockouts (SPK), dual parameter knockouts (DPK) and multiple parameter knockouts (MPK) to the EHR network. In the SPK study, we found that on average 200 parameters led to increased marker levels, while around 350 parameters led to decreased marker levels. Deletion of interactions involved with PHD formation, HIF-1α degradation, the regulation of decoy receptors for VEGFA and the sequestering of VEGFA and IL-8 signals resulted in the most pronounced effects. This result supports the claim of the structural analysis that VEGF and IL-8 signal integration and HIF-1α regulation in hypoxia are the most critical aspects of EHR. SPK analysis also highlighted the importance of AP-1, p38MAPK and PKC. This result is in accord with the positive feedback mediated by these species on the EHR, so any alterations in the regulation of AP-1, p38MAPK and PKC have a reduction in the marker levels. Phenotypic behavior of the EHR models were analyzed as a result of parameter knockout, using downstream markers like BNIP3 (marker for cell death), Bcl2 (marker for cell survival) and Cyclin D (marker for cell growth) (supplementary materials Fig. A.4). We identified 35% of SPK cases which had a pro-death phenotype marked by increased (robustness coefficient > 1) marker levels of BNIP3 and decreased (robustness coefficient < 1) marker levels of Bcl2 (supplementary materials). Around 20% of SPK cases had a pro-growth phenotype marked by increased (robustness coefficient > 1) marker levels of Cyclin D as compared to pro-death phenotype (supplementary materials). Results clearly shows the *+ve* correlation between cell growth and cell survival phenotypic behavior. It is interesting to note the feasibility of regulating the phenotypic behavior by altering one or more interactions. This highlights the importance of a structurally complete model in order to analyze the overall behavior.

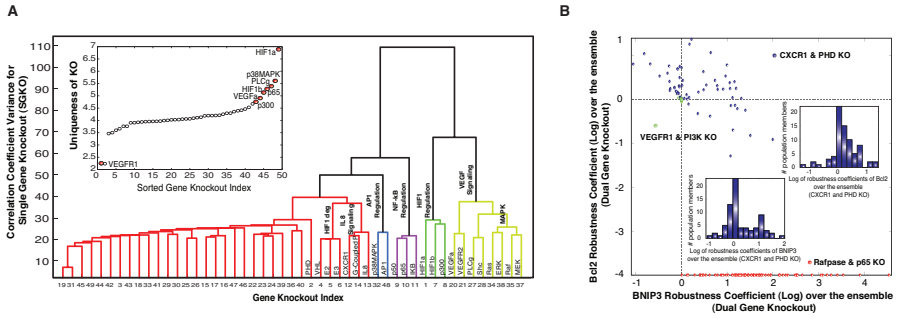


Figure 6.8: Robustness Analysis of EHR models: (A) Dendrogram of the single gene knockouts revealed clusters with inherent functional relationships. For example, the most distinct separation was between the MAPK, PKC and VEGF signal initiation segment from the rest of the knockouts. MAPK and PKC are one of the few pillars mediating the crosstalk between VEGF and IL-8 signaling and also regulating signal amplification. On the other hand, functional groups like regulation of NF-$\kappa\beta$, regulation of AP-1, HIF-1 regulation and IL-8 signal initiation were all rightly classified as having similar grouped responses. Inset - A: Plot of the orthogonal components of each knockout so as to identify the uniqueness of the knockout. All the single gene knockouts were found to have orthogonal components greater than one with a 95% confidence. Knockout of HIF-1α, p38MAPK, p65, PLCγ and VEGFA were some of the knockouts which produced the most unique knockouts. (B) Phenotypic analysis (pro-death as elevated BNIP3 levels or pro-survival as elevated Bcl2 levels) of dual gene knockouts using three cases: case 1 - CXCR1 & PHD KO, case 2 - VEGFR1 & PI3K KO and case 3 - RAFPase & p65 KO. In case 1, there was a diverse distribution of response in terms of pro-death phenotype (marked by log of robustness coefficients of BNIP3 both > and < 1). We found two populations, one set of ensembles (~25%) revealed a pro-death phenotype marked by a ten-fold increment in BNIP3 levels. On the other hand, there was a population of the ensemble (~50%) which did not have much effect over the BNIP3 levels. This attested the fact that we have a diversity in terms of cellular behavior in the ensemble of our EHR models. In case 3, RAFPase and p65 KO lead to a dominant pro-death phenotype.

Single gene knockdown and over-expression studies are common in the experimental literature. However, large-scale single or combinatorial knockdown experiments similar to, for example the Keio collection in *E. coli* [20], are not tractable in human or mouse hosts. Comprehensive knockout and over-expression simulations, using validated mechanistic models, are perhaps an alternative means to generate testable hypotheses and therapeutic insights. We looked at 548 markers following the deletion or overexpression of the 49-regulated genes from the model (supplementary material Fig. A.5). Gene knockout studies revealed the critical players involved in increase of marker levels as components of the HIF-1α degradation machinery (e.g., PHD) and HIF-1α regulation (e.g., FIH), decoy receptors for the VEGFA (e.g., VEGFR1) and regulatory elements of the MAPK cascades (e.g., ERKPase) (supplementary materials Fig. A.5). However, the predominant effect of gene knockouts was that of reduced marker levels. Early stage markers of the EHR signaling cascade like HIF-1α, HIF-1β, VEGFA, VEGFR2, PKC, p38MAPK, NF-$\kappa\beta$ and AP-1 were clearly seen to have an enhanced down-regulatory effect (Fig. A.5A). Global analysis of the gene knockout robustness coefficients by clustering to minimize the variance, provides a systems-level insight into EHR. As in Fig. **??**A, dendrogram of the single gene knockouts revealed clusters with inherent functional relationships. For example, the most distinct separation was between the MAPK, PKC and VEGF signal initiation segment from the rest of the knockouts. MAPK and PKC are one of the few pillars mediating the crosstalk between VEGF and IL-8 signaling and also regulating signal amplification. On the other hand, functional groups like regulation of NF-$\kappa\beta$, regulation of AP-1, HIF-1 regulation and IL-8 signal initiation were all rightly classified as having similar grouped responses. We computed the magnitude of the orthogonal components

of the single gene knockout coefficients (Fig. 6.8A Inset). The orthogonal component was used to establish the uniqueness of the knockout. All the single gene knockouts were found to have orthogonal components greater than one with a 95% confidence. Knockout of HIF-1α, p38MAPK, p65, PLCγ and VEGFA were some of the knockouts which produced the most unique knockouts. Interestingly, knockouts implicating or regulating the transcription factors are the most influential knockouts. Gene overexpression results were complementary to the knockout studies, revealing the key role of regulation of HIF-1α signal initiation, regular and decoy VEGFA signaling and components of MAPK (MEK, ERK) (Fig. A.5B). Complementary to normal intuitions, increase in HIF-1α or IL-8 alone did not lead to up-regulation of the EHR due to crosstalk mediated by NF-$\kappa\beta$ and AP-1 in regulating VEGF and IL-8 signaling cascades. However, overexpression of CXCR1, VEGFA, VEGFR2, ERK and MEK lead to increased marker levels and enhanced EHR (Fig. A.5B). Dendrogram of the clusters of overexpression studies revealed that IL-8, AP-1 regulation and HIF-1 regulation had the most distinct separation from the rest of the overexpression studies (Fig. A.7). Around 96% of the single gene overexpression studies were found to have orthogonal components greater than one with a 95% confidence (Fig. A.7 Inset).

Using combinatorial gene knockouts (1176 exclusive combinations of two genes), we explored the phenotypic behavior exhibited by the EHR model (Fig. A.6). Combinatorial gene knockouts identified unique phenotypes as compared to the phenotypic behavior obtained by SPK-DPK analysis. We identified sub-populations that had a pro-survival (15%), pro-growth (15%) and pro-death (65%) phenotype. The positive correlation between pro-growth and pro-survival phenotypes could be attributed to the transcriptional role of AP-1 and

NF-$\kappa\beta$. The dominant phenotype of pro-death could be attributed to the HIF-1 dependent regulation of BNIP3 and NF-$\kappa\beta$ dependent regulation of Bcl2. However the dominant phenotype in tumors and cancer is that of pro-growth and pro-survival. It was interesting to note populations that were both pro-death and pro-survival. Synergistic effects in combinatorial gene knockouts were due to the crosstalk of VEGF and IL-8 signaling cascades mediated by the role of NF-$\kappa\beta$ and AP-1.

We isolated three cases to investigate the dual gene knockouts further: case 1 - CXCR1 & PHD KO, case 2 - VEGFR1 & PI3K KO and case 3 - RAFPase & p65 KO and analyzed their effect over the entire ensemble (Fig. 6.8B). We analyzed the effects of these KOs in the pro-growth and pro-death phenotypic plane. In case 1, there was a diverse distribution of response in terms of pro-death phenotype (marked by log of robustness coefficients of BNIP3 both > and < 1) (Fig. 6.8B Inset). We found two populations, one set of ensembles (~25%) revealed a pro-death phenotype marked by a ten-fold increment in BNIP3 levels. On the other hand, there was a population of the ensemble (~50%) which did not have much effect over the BNIP3 levels. This attested the fact that we have a diversity in terms of cellular behavior in the ensemble of our EHR models. BNIP3 activity is regulated both at transcriptional level via HIF-1 and at the protein level via binding to Bcl2. PHD knockout, leads to stronger HIF-1 signal so at the transcriptional level we have more BNIP3 production. However, the role of Bcl2 is regulated via NF-$\kappa\beta$ which is regulated by both HIF-1 dependent and HIF-1 independent cascades. So the difference in the BNIP3 levels in the population, could be attributed to the differential regulation of NF-$\kappa\beta$ in the ensemble. In case 2, we did not find much effect of the KO over the ensemble. This was attributed to the lack of structural connectivity between active PI3K and

the other components of the model. Active PI3K was a downstream component in the current EHR model and was only connected to Paxillin. VEGFR1 in the model only acts as a decoy receptor for VEGFA, so knockout of VEGFR1 only leads to amplification of the hypoxic signal in the VEGF branch. However in recent experiments it was shown that VEGFR1 signals through the eNOS pathway leading to implications in the VEGF signaling pathway [535], which has not been currently included in the model. In case 3, RAFPase and p65 KO lead to a dominant pro-death phenotype. This behavior was captured by the model by the lack of NF-$\kappa\beta$ activity leading to no Bcl2 production and enhanced BNIP3 activity. This observation was consistent with the literature, where in RelA or p65 KO lead to embryonic lethality and liver degeneration in mice [30]. To successfully manipulate the hypoxic response, it is therefore vital to regulate both NF-$\kappa\beta$ and AP-1 regulatory signaling blocks of EHR. The current EHR model, even with inherent structural and parametric uncertainty, identifies falsifiable strategies in manipulating the hypoxic response.

6.4 Discussion

Eukaryotic hypoxic response (EHR) blankets a diverse set of responses mediated by a single physiological stimulus i.e. lack of sufficient oxygen. EHR globally leads to altered cellular functioning, stimulation of erythropoiesis, angiogenesis, vascular tone regulation, altered cell viability and proliferation [12]. The activation of the angiogenic switch represents a critical event in the vascularization of primary and metastatic tumors and thus in the progression of the pathology. The underlying cellular and molecular mechanisms controlling the angiogenic switch are still emerging. However, what is generally accepted is

that changes in the tumor microenvironment e.g., extracellular oxygen concentration, induce a multifaceted program involving both autocrine and paracrine signaling amongst an array of cell types. Specifically, the role of EHR in angiogenesis has become an area of interest because of its potential as a therapeutic locus [349, 50]. In this study, we formulated a family of canonical EHR models describing events in the initial phase of the angiogenic switch. Mass balance equations describing 548 species interconnected by 920 interactions were formulated using mass-action kinetics within an ordinary differential equation (ODE) framework. Four model populations were estimated using multiobjective optimization (20 objective functions) in conjunction with a leave-five-out cross-validation strategy. We compared the average performance of the models in the population, by comparing training and prediction likelihood. The multiobjective optimization scheme (POETs) improved the training and prediction likelihood for 70% of the objective functions, relative to an initial best-fit parameter set. POETs also identified Pareto fronts between a subset of the training objectives. These fronts indicated conflicts in the training data e.g., cell-line to cell-line variation or an inability to simultaneously model different aspects of the training data. POETs generated a population of 40,000 probable models, from which we selected a diverse sub-family with Pareto rank one or less for further study. To identify critical nodes, the EHR network and model populated were analyzed using parameter dependent and independent analysis techniques. Model analysis based on the flux allowed a temporal analysis of EHR. Parameter-dependent sensitivity and robustness analysis were used to rank-order the importance of nodes and edges in the EHR network. We also analyzed the topological features of the EHR network using extreme pathway decomposition, in particular identifying the structural basis of crosstalk and re-

dundancy.

EHR model analysis revealed that there is a temporal progression within EHR, which has its implications in terms of parameter(species) robustness(sensitivity) and correspondingly affects complex network properties like crosstalk and redundancy within EHR. Under hypoxic conditions (~ 1% O_2), we found that there was a two-phased response demarcating the onset of a later autocrine mediated signal amplification. While initially both HIF-1 and ROS mediated activation of VEGF and IL-8 signaling cascades took place, there was a late phase amplification of VEGF and IL-8 consistent with autocrine mediated signal amplification due to AP-1. This result was validated by independent ELISA measurements of secreted VEGF and IL-8 in MDA-MB231 cells at days 1,3 and 5. In phase I, VEGF signaling branch was the first to respond, mediated by the activity of HIF-1 and ROS mediated NF-$\kappa\beta$ activation. However, VEGF levels reached steady state by 24 hrs (seen both by simulation and experimental study). Phase II of the response was marked by a small increment of VEGF and 3-5 fold increment in IL-8 signaling. This late phase behavior was a result of the positive feedback mediated by the activation of AP-1 and NF-$\kappa\beta$ downstream of both VEGF and IL-8 signaling cascades. This phenomenon is consistent with previous studies in OSCC-3 cells in both *in-vivo* and *in-vitro* studies [129]. However, the time scales of the signal amplification was different between model prediction (~ 24 hrs) and as seen in the experiments (~ 48 hrs). This difference could be attributed to the negative feedback mediated by HIF-1 induction which leads to attenuated Nrf-2 dependent IL-8 expression as seen in HUVECs [315]. Lack of this negative control on IL-8 expression via HIF-1, could be one of the many reasons for the high predicted rate of IL-8 secretion in the model. Interestingly, it was observed both experimentally and predicted by the

model that under normoxic conditions secreted IL-8 levels increased with time. One potential explanation could be that, normoxic conditions lead to even lower amount of basal HIF-1 activation and subsequently more Nrf-2 mediated IL-8 expression. Overall, the population of EHR models recapitulated both short- and long timescale behavior following hypoxia or exposure to exogenous IL-8 in a variety of cell types (training and prediction studies). It also captured the integration between multiple pathways and generated specific and testable hypothesis about the role of network components in signal propagation.

Consistent with the key regulatory role of AP-1 and NF-$\kappa\beta$ in EHR, we identified that in case of 1% O_2 (hypoxia), MAPK regulation mediated by MEKPase (phosphatase for MEK) and VEGFR2-mediated signaling was the most important in terms of participation in EPs (~300 EPs). On the contrary, in case of 17% O_2 (normoxia), HIF-1α regulation via O_2 was the set of reactions with the most participation (~170 EPs). In case of complex biological networks like EHR, interactions between specific or shared species between multiple pathways gives rise to intricate complexities within the network. These complexities critically affect the dynamics and properties of the system. Currently research techniques aim at identification of global network properties [234, 459, 277]. Signal transduction architectures however, frequently contain redundancy, feedback and crosstalk. These topological features ensure signal propagation is adaptable, efficient and robust. However, they also make reprogramming signal flow challenging. Extreme pathway decomposition (and the closely related elementary mode decomposition) systematically decompose large networks into independent mass-balanced pathways [482, 406]. Thus, instead of a relying on arbitrary pathway definitions found in the literature, EPs could be used to systematically decompose complex networks. EPs can also be used to identify the structural

basis of complex network properties such as crosstalk and redundancy. In this study, EPs for EHR provided a structural significance of key components in any mode of functioning of EHR. The structural analysis corroborated the structural relevance of MAPK in hypoxic conditions. Using EPs, we found that in case of hypoxia ~60% of the EPs exhibited crosstalk, while ~5% of the EPs exhibited redundancy. Overall, EP analysis of EHR suggested a divergent nature of the EHR. This suggests that, oxygen sensing lead to HIF-1 and ROS mediated activation of VEGF and IL-8 signals, subsequently leading to multiple outputs like active PI3K, active MAPK components, active transcription factors e.g. NF-$\kappa\beta$, AP-1 which have the capability to communicate with other signaling modules in the cell. The redundancy within EHR, reiterates multiple routes of activation and feedback of key transcription factors like AP-1, NF-$\kappa\beta$.

Consistent with the later time importance of AP-1, regulators of AP-1 i.e. p38MAPK and MAPK were seen as fragile aspects at ~ 24-30 hrs into Hypoxia as compared to earlier time points in the sensitivity analysis. Sensitivity analysis is an enabling tool for the investigation of fragility in molecular networks. It has been used earlier to identify fragile interactions in cell-cycle models [369] and in human coagulation cascades [319] as well in other systems [274]. It has been suggested that sensitivity analysis of experimentally constrained model ensembles could in fact generate a reasonable estimate of what was important in a network [504]. In terms of species sensitivity in the case of 1% O_2, MAPK components were seen to be more important at later (24-30 hrs) time band as compared to earlier (0-6 hrs) times of EHR. This highlights the critical role of MAPK as a critical regulator of AP-1 activation in autocrine mediated signal amplification. Both parameter and species sensitivities suggested that AP-1 and NF-$\kappa\beta$ regulatory module gained importance as time progressed

into EHR. This emphasizes the regulatory role played by AP-1 and NF-$\kappa\beta$ in influencing the positive feedback downstream of both VEGF and IL-8 signaling cascades. Interestingly, both VEGF and IL-8 signaling modules were relatively equally fragile(robust) throughout EHR. While sensitivity analysis identifies the critical elements in the EHR network, it does not elucidate the effect of a specific perturbation on the other components in the system. In this regard, simulated gene-knockout (single/combinatorial) and gene over-expression experiments for the EHR model using robustness analysis, allowed us to identify critical components to manipulate the hypoxic response. Robustness analysis has been used previously in molecular networks for androgen dependent and independent prostate cancer models to identify fragile and robust subsystems [526]. This simulation scheme bears resemblance to current experimental protocols, where we regulate the hypoxic programs are via gene knockouts or gene over-expression. Up-regulation of the hypoxic program has been used successfully in studies related to transplantation, wherein angiogenesis is a desired and critical aspect for survival of cells [400, 604]. On the contrary, in cancer therapy down-regulation of hypoxic response is a much sought after target [349]. Gene knockout and two-fold overexpression studies revealed the critical players as components of the HIF-1α degradation machinery (e.g., PHD) and HIF-1α regulation (e.g., FIH), decoy receptors for the VEGFA (e.g., VEGFR1) and regulatory elements of the MAPK cascades (e.g., ERKPase). In the phenotypic space as a result of single/combinatorial gene knockout/overexpression simulations, we identified sub-populations that had distinct pro-survival, pro-growth and pro-death phenotypes. Behavioral shifts in the phenotypic space (like positive correlation between pro-growth and pro-survival phenotypes) were due to the role of AP-1 and NF-$\kappa\beta$. The dominant phenotype of pro-death could be at-

tributed to the HIF-1 dependent regulation of BNIP3 and NF-$\kappa\beta$ dependent regulation of Bcl2. Synergistic effects controlled by crosstalk and redundancy in the EHR network thereby controls the phenotypic behavior of the EHR models upon manipulation. In order to successfully manipulate the hypoxic response, it is therefore vital to regulate both NF-$\kappa\beta$ and AP-1. This highlights the importance of regulating both HIF-1 and ROS mediated signaling blocks. Recent experimental studies have shown combination strategies aimed at maximizing the clinical potential of HIF-1 inhibition, to be highly effective [441].

Overall, model analysis suggested relationships between structure, function and regulation within the EHR network and provide testable hypothesis which can be validated experimentally. Key aspects of the EHR model were centered around regulation and manipulation of HIF-1α, AP-1 and NF-$\kappa\beta$. Importance of HIF-1α regulation highlighted by the structural and robustness analysis is supported by the evidence in literature wherein inhibition of HIF-1α is used as a strategy in tumor therapy [552, 484]. However we are lacking some details in regulation of HIF-1α, which taken into account might reveal a clearer picture [25]. In the current EHR model, HIF-1α was regulated only at the protein level. We only have the constitutive expression of HIF-1α and HIF-1β. In order to get a complete picture of regulation of HIF-1α, we should include the negative regulation of HIF-1α via the regulation of PHD2 protein. PHD2 mRNA and protein have been reported to be upregulated in case of hypoxia, leading to negative feedback [25]. Transcriptional role of HIF-1 is mediated via binding to p300 protein. CITED2/4 proteins (increased expression in hypoxia) have been identified which competitively binds to p300 thereby operating in a negative feedback loop to down-regulate HIF-1 mediated transactivation [25]. p53 has been seen to promote Mdm2-mediated ubiquitination and proteasomal degra-

dation of HIF-1α [445]. Furthermore there exists a positive feedback to HIF-1α signaling mediated by growth factor and cytokines via Ras/Raf/MAPK and PI3-K/Akt/PKB cascades [25]. These pathways also play a role in regulating the HIF-1 signaling cascades in Normoxia. Given the role of VEGF and IL-8 signaling pathways in angiogenesis, it has rightly been targeted in several efforts in regulating tumors [591, 271, 126]. Regulation of VEGF by the humanized monoclonal antibody bevacizumab, is currently used successfully for the treatment of cancer [127]. In the current model, VEGF mediated signal initiation has been modeled via VEGFA, VEGFR1/2/3, heparin sulfate proteoglycans (hspg) and neuropilins. Given the importance of VEGF signaling, we might get a clearer and more specific picture upon inclusion of specific details regarding variants of VEGFA (VEGF-206, VEGF-165 and VEGF-183), other ligands like VEGFB/C/D/E and the specific role of Neuropilin-1 [458]. Neuropilin-1 mRNA and protein have been seen to be overexpressed in case of human colon growth and progression [409]. Downstream effects of VEGF signaling cascade mediated by PI3K-AKT module have also not been taken into account in the current EHR model [562, 69]. Both structurally and functionally relevant, role of AP-1 and NF-$\kappa\beta$ has been identified by the model analysis to be of high importance. Rightfully so, regulation of NF-$\kappa\beta$ has been seen to be critically involved in cancer development and progression, as well as in resistance to chemotherapy and radiotherapy [27]. Thereby, the therapeutic potential and benefit of targeting NF-$\kappa\beta$ and ROS in cancer holds great promise [27, 555]. Currently, a narrow account of NF-$\kappa\beta$ regulation has been taken into account in the EHR model. It shall be interesting to incorporate the detailed complexity associated with regulation of NF-$\kappa\beta$ and then analyze the critical aspects of the EHR model [86]. In terms of NF-kB, we have not taken into account the regulation of IKK

via PKC. IKK traditionally leads to Ser32 and Ser36 phosphorylation of IKBα, thereby targeting it for proteosomal degradation [86]. However in case of hypoxia another alternate route of IKBα phosphorylation not including proteosomal degradation also plays a key part [86]. In the current EHR model, we have AP-1 regulation via p38MAPK and ROS. We have not taken into account the individual dimer components, namely cFOS and cJUN. To get a comprehensive insight into the role of AP-1, we should include specific details like ERK2 mediated phosphorylation of cFOS at T232 to establish the link between ERK and AP-1 [23]. MAP kinases ERK, JNK and p38 phosphorylate and activate Elk-1 resulting in enhanced SRE-dependent cFOS expression [586]. JNK and p38 MAP kinases lead to activation and enhanced expression of cJUN [586]. Given structural and parametric uncertainty in the given scheme of model generation, the EHR models were successful in identifying the critical components of EHR and present valuable insight in regulating the hypoxic program. Structural analysis coupled with concurrent developments in systems biology provides a scope for a functional and mechanistic basis for design of simulated experiments and prediction of critical targets in large-scale complex biological models.

6.5 Materials and Methods

6.5.1 Formulation and solution of the model equations.

The eukaryotic hypoxic response model was formulated as a set of coupled ordinary differential equations (ODEs):

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{r}(\mathbf{x}, \mathbf{p}) \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{6.1}$$

The symbol **S** denotes the stoichiometric matrix (548 × 920). The quantity **x** denotes the concentration vector of proteins or protein complexes (548 × 1). The term $\mathbf{r}(\mathbf{x}, \mathbf{p})$ denotes the vector of reaction rates (920 × 1). Each row in **S** described a protein or protein-protein complex, while each column described the stoichiometry of network interactions. Thus, the (i, j) element of **S**, denoted by σ_{ij}, described how protein i was involved in rate j. If $\sigma_{ij} < 0$, then protein i was consumed in r_j. Conversely, if $\sigma_{ij} > 0$, protein i was produced by r_j. Lastly, if $\sigma_{ij} = 0$, there was no protein i in rate j. All of these interactions were obtained from the literature.

We assumed mass-action kinetics for each interaction in the network. The rate expression for interaction q was given by:

$$r_q\left(\mathbf{x}, k_q\right) = k_q \prod_{j \in \{\mathbf{R}_q\}} x_j^{-\sigma_{jq}} \tag{6.2}$$

The set $\left\{\mathbf{R}_q\right\}$ denotes reactants for reaction q while σ_{jq} denotes the stoichiometric coefficient (element of the matrix **S**) governing species j in reaction q. All reversible interactions were split into two irreversible steps. The mass-action formulation, while expanding the dimension of the EHR model, regularized the mathematical structure; this allowed automatic generation of the model code using UNIVERSAL and regularized the unknown model parameters (parameters were one of only three types, association, dissociation or catalytic rate constants). UNIVERSAL, an open source Java code generator, generates multiple code types from text and SBML inputs. UNIVERSAL is freely available as a Google Code project (http://code.google.com/p/universal-code-generator/). Thus, although mass-action kinetics increased the number of parameters and species, they reduced the complexity of model analysis. In this study, we considered well-mixed nuclear, cytosolic and extracellular compart-

ments. The model equations were solved using the LSODE routine in OCTAVE (v 3.1.0; www.octave.org) on an Apple workstation (Apple, Cupertino, CA; OS X v10.6.4).

Hypoxic conditions were simulated by running the model to steady state and then reducing the oxygen input. The steady-state was estimated numerically by repeatedly solving the model equations and estimating the difference between subsequent time points:

$$\|\mathbf{x}(t+\Delta t)-\mathbf{x}(t)\|_2 \leq \gamma \tag{6.3}$$

The quantities $\mathbf{x}(t)$ and $\mathbf{x}(t+\Delta t)$ denote the simulated concentration vector at time t and $t+\Delta t$, respectively. The L_2 vector-norm was used as the distance metric. We used $\Delta t = 1$ s and $\gamma = 0.001$ for all simulations.

6.5.2 Estimation and cross-validation of a population of models using Pareto Optimal Ensemble Techniques (POETs).

POETs is a multiobjective optimization strategy which integrates several local search strategies e.g., Simulated Annealing (SA) or Pattern Search (PS) with a Pareto-rank-based fitness assignment [504]. Denote a candidate parameter set at iteration $i+1$ as $\mathbf{k}_{i+1}$. The squared error for $\mathbf{k}_{i+1}$ for training set j was defined as:

$$E_j(\mathbf{k}) = \sum_{i=1}^{\mathcal{T}_j}\left(\hat{\mathcal{M}}_{ij} - \hat{y}_{ij}(\mathbf{k})\right)^2 \tag{6.4}$$

The symbol $\hat{\mathcal{M}}_{ij}$ denotes scaled experimental observations (from training set j) while the symbol $\hat{y}_{ij}$ denotes the scaled simulation output (from training set j). The quantity i denotes the sampled time-index and $\mathcal{T}_j$ denotes the number of

time points for experiment j. The read-out from the training immunoblots was band intensity where we assumed intensity was only loosely proportional to concentration. Suppose we have the intensity for species x at time $i = \{t_1, t_2, .., t_n\}$ in condition j. The scaled measurement would then be given by:

$$\hat{\mathcal{M}}_{ij} = \frac{\mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}}{\max_i \mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}} \tag{6.5}$$

Under this scaling, the lowest intensity band equaled zero while the highest intensity band equaled one. A similar scaling was defined for the simulation output.

We computed the Pareto rank of $\mathbf{k}_{i+1}$ by comparing the simulation error at iteration $i+1$ against the simulation archive $\mathbf{K}_i$. We used the Fonseca and Fleming ranking scheme [133]:

$$rank\left(\mathbf{k}_{i+1} \mid \mathbf{K}_i\right) = p \tag{6.6}$$

where p denotes the number of parameter sets that dominate parameter set $\mathbf{k}_{i+1}$. Parameter sets on or near the optimal trade-off surface have small rank. Sets with increasing rank are progressively further away from the optimal trade-off surface. The parameter set $\mathbf{k}_{i+1}$ was accepted or rejected by the SA with probability $\mathcal{P}\left(\mathbf{k}_{i+1}\right)$:

$$\mathcal{P}(\mathbf{k}_{i+1}) \equiv \exp\left\{-rank\left(\mathbf{k}_{i+1} \mid \mathbf{K}_i\right)/T\right\} \tag{6.7}$$

where T is the computational annealing temperature. The initial temperature $T_o = n/log(2)$, where n is user defined ($n = 4$ for this study). The final temperature was $T_f = 0.1$. The annealing temperature was discretized into 10 quanta between T_o and T_f and adjusted according to the schedule $T_k = \beta^k T_0$ where β was defined as:

$$\beta = \left(\frac{T_f}{T_o}\right)^{1/10} \tag{6.8}$$

The epoch-counter k was incremented after the addition of 100 members to the ensemble. Thus, as the ensemble grew, the likelihood of accepting parameter sets with a large Pareto rank decreased. To generate parameter diversity, we randomly perturbed each parameter by $\leq \pm 25\%$. We performed a local pattern-search every q steps to minimize the residual for a single randomly selected objective. The local pattern-search algorithm has been described previously [149, 560]. The parameter ensemble used in the simulation and sensitivity studies was generated from the low-rank parameter sets in $\mathbf{K}_i$.

We simultaneously calculated training and prediction error during the parameter estimation procedure using leave-five-out cross-validation [261]. The complete set of training data (20 objectives) was subdivided into four bins; in each bin 15 data sets were reserved for training while five were reserved for prediction. In the first bin $DS_1 \ldots DS_5$ were used for validation while $DS_6 \ldots DS_{20}$ were used for training. In the second bin $DS_6 \ldots DS_{10}$ were used for validation while $DS_1 \ldots DS_5$ $DS_{11} \ldots DS_{20}$ were used for training, etc. Thus, we formulated four ensembles from which we evenly selected parameter sets for the *parent* ensemble (Fig. A.1). While cross-validation required that we generate additional model populations, we trained and tested against all the data sets.

6.5.3 Sensitivity and robustness analysis of the population of EHR models.

Sensitivity coefficients were calculated as shown previously [505, 504, 526] using 40 models selected from the ensemble (red points, supplementary materials Fig. A.1). The resulting sensitivity coefficients were scaled and time-averaged

(Trapezoid rule):

$$\mathcal{N}_{ij} \equiv \frac{1}{T} \int_0^T dt \cdot |\alpha_{ij}(t)\, s_{ij}(t)| \tag{6.9}$$

where T denotes the final simulation time and $\alpha_{ij} = 1$. The time-averaged sensitivity coefficients were then organized into an array for each ensemble member:

$$\mathcal{N}^{(\epsilon)} = \begin{pmatrix} \mathcal{N}_{11}^{(\epsilon)} & \mathcal{N}_{12}^{(\epsilon)} & \dots & \mathcal{N}_{1j}^{(\epsilon)} & \dots & \mathcal{N}_{1P}^{(\epsilon)} \\ \mathcal{N}_{21}^{(\epsilon)} & \mathcal{N}_{22}^{(\epsilon)} & \dots & \mathcal{N}_{2j}^{(\epsilon)} & \dots & \mathcal{N}_{2P}^{(\epsilon)} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathcal{N}_{M1}^{(\epsilon)} & \mathcal{N}_{M2}^{(\epsilon)} & \dots & \mathcal{N}_{Mj}^{(\epsilon)} & \dots & \mathcal{N}_{MP}^{(\epsilon)} \end{pmatrix} \qquad \epsilon = 1, 2, \dots, N_\epsilon \tag{6.10}$$

where ϵ denotes the index of the ensemble member, P denotes the number of parameters, N_ϵ denotes the number of ensemble samples and M denotes the number of model species. To estimate the relative fragility or robustness of species and reactions in the network, we decomposed the $\mathcal{N}^{(\epsilon)}$ matrix using Singular Value Decomposition (SVD):

$$\mathcal{N}^{(\epsilon)} = \mathbf{U}^{(\epsilon)} \Sigma^{(\epsilon)} \mathbf{V}^{T,(\epsilon)} \tag{6.11}$$

Coefficients of the left (right) singular vectors corresponding to largest β singular values of $\mathcal{N}^{(\epsilon)}$ were rank-ordered to estimate important species (reaction) combinations. Only coefficients with magnitude greater than a threshold (δ = 0.1) were considered. The fraction of the β vectors in which a reaction or species index occurred was used to rank its importance.

Robustness coefficients of the form:

$$\alpha\left(i, j, t_o, t_f\right) = \left(\int_{t_o}^{t_f} x_i(t)\, dt\right)^{-1} \left(\int_{t_o}^{t_f} x_i^{(j)}(t)\, dt\right) \tag{6.12}$$

were calculated to understand the robustness of the network. The robustness coefficient $\alpha\left(i, j, t_o, t_f\right)$ is the ratio of the integrated concentration of a network marker in the presence (numerator) and absence (denominator) of structural

or operational perturbation. The quantities t_0 and t_f denote the initial and final simulation time respectively, while i and j denote the indices for the marker and the perturbation respectively. If $\alpha\left(i, j, t_o, t_f\right) > 1$, then the perturbation *increased* the marker concentration. Conversely, if $\alpha\left(i, j, t_o, t_f\right) \ll 1$ the perturbation *decreased* the marker concentration. Lastly, if $\alpha\left(i, j, t_o, t_f\right) \sim 1$ the perturbation did not influence the marker concentration.

In this study, we simulated single and double edge knockouts along with protein deletion and overexpression. For multiple simultaneous perturbations, combinatorial effects were evaluated by calculating synergy coefficients. The synergy coefficient for a marker k following a combinatorial perturbation j, denoted by $\mathcal{S}_{ij}$, was

$$\mathcal{S}_k^{i,j} = \frac{\alpha_k^{(i,j)}}{\alpha_k^i + \alpha_k^j} \qquad i, j = \left(1, 2, \ldots, N_{pk}\right) \tag{6.13}$$

defined as the robustness coefficients for marker i in the presence of multiple simultaneous perturbations (set j) normalized by the sum of the individual robustness coefficients of each single perturbation in set j.

6.5.4 Species clustering and dendrogram.

A dendrogram was derived by considering each of the knockouts(overexpressions) as variables and the average log of robustness coefficient (LRC) for each of the species as observations. We used the Euclidean norm in LRC space as the distance metric. The linkage function (objective function for identifying variable clusters) was the inner squared distance (minimum variance algorithm). The Statistical Toolbox of Matlab (The Mathworks, Natick, MA) was used to generate the distances, linkages and the final dendrogram.

6.5.5 Identification of distinguishable species.

Robustness coefficients were used to rank-order knockout(overexpression) experiments in terms of the greatest unique responses and identify species which were linearly distinguishable. The response of the knockout(overexpression) was measured in terms of the robustness coefficients. The LRC had desirable linear properties, such that no response (no change in trajectories from wild-type) returns a value of zero and similar negative and positive responses have different directions but similar magnitudes. We considered the unique component of the response to be the orthogonal component in LRC space and the magnitude of the response to be the Euclidean norm. The orthogonal components and there magnitude were identified for each parameter set in the ensemble by first choosing the knockout(overexpression) with the greatest magnitude, x_1 and placing it in the empty set $\mathcal{V}$. The knockout(overexpression) x_1 defines the orthogonal directions in the LRC space. We then calculated the orthogonal components for all remaining knockouts(overexpressions) relative to x_1, and added the knockout(overexpression) species with the greatest orthogonal magnitude to set $\mathcal{V}$. In general the components of all remaining x_i orthogonal to set $\mathcal{V}$ were calculated and the largest was moved into set $\mathcal{V}$. This process was continued until all knockout(overexpression) species, x_i were added to set $\mathcal{V}$. Mathematically two species were considered distinguishable if and only if they were linearly independent (the orthogonal components were non-zero). We considered a threshold value of one or five and performed a student t-test (Matlab Statistical Toolbox, The Mathworks, Natick, MA) to identify which species had orthogonal components above the threshold with a 95% confidence over the ensemble.

6.5.6 3-D Alginate Tumor Models

Alginate hydrogels (3% [w/v]) were used as 3-D cell scaffolds. Alginate (Protanal LF; FMC Biopolymer, Philadelphia, PA) was dissolved in serum-free DMEM, and MDA-MB231 cells (ATCC) were suspended in this dissolved alginate at a density of 20×10^6 cell/mL. Cell-seeded alginate gels were cast into 4 mm diameter, 200 μm thick cylinders using a plexiglass mold, and cross-linked with 100 mM $CaCl_2$ for 15 min. Discs were then cultured in 24-well plates, with a single disc and 500 μL culture media (DMEM + 10% FBS + 1% penicillin/streptomycin) per well. Cells were cultured at 37°C, 5% CO_2 and either 1 or 17% O_2 in controlled atmosphere incubators (Thermo Fisher Scientific, Inc., Waltham, MA). Media was changed daily.

6.5.7 Analysis of Angiogenic Factor Secretion

Media and discs were collected 24 hrs after the previous media change, at days 1, 3 and 5. VEGF and IL-8 concentration in media were analyzed by ELISA (R&D Systems, Minneapolis, MN) and normalized to total DNA content to correct for cell number variations. To measure DNA content, discs were dissolved in 50 mM EDTA and released and rinsed cells lysed in Caronś buffer. DNA concentration was then measured using Quant-iT PicoGreen dsDNA reagent (Invitrogen Co.).

CHAPTER 7

SIGNALING ASSISTED MULTISCALE MODELING (SAMM) OF BREAST CANCER PROGRESSION

Authors - Chakrabarti A, Chan JF, Lequieu JP, Paull M, Balaji C, Tasseff RA, Verbridge SS, Stroock AD, Fischbach-Teschl C and Varner JD

Submitted to - Annals of Biomedical Engineering

7.1 Abstract

Several disease states like cancer are heterogeneous cellular entities involving dynamical interactions between genetically altered cells and their surrounding microenvironmental changes. These forms of disease states present an emerging importance and need of models taking both signaling networks and different biochemical and biophysical contexts into account. For example, factors like vascular endothelial growth factor (VEGF) and/or interleukin-8 (IL-8) have been implicated in multiple studies in the regulation of breast cancer as well as in the activation and chemotactic guidance of capillary sprouting and growth. Detailed computational models of such complex phenomena using intricate cellular signaling mechanisms, nutrient/growth factor transport, movement/interactions between normal and tumor cells is critical in identifying the key regulators of such a system. We have created a protein signaling assisted multi-scale modeling approach (SAMM) to study such systems. Key highlights of SAMM include use of pre-established detailed cellular models to regulate cellular decisions, use of an ensemble of models to incorporate differences between cells, coupling intracellular markers and probabilistic rules to monitor

cellular fate and cellular motility and a hybrid approach to improve speed of solving of external and internal mass-balances. SAMM allows us the flexibility to include multiple cell types, freedom of seeding cells with specific internal cellular machinery within complex geometries, regulate multiple cellular decisions using internal and secreted cellular markers and improved solving speed. In the current study, we demonstrated using SAMM and a tumor/endothelial cell model the ability to capture and identify spatio-temporal effects upon GF (VEGF and/or IL-8) induced growth, ability to model and monitor effects of internal variations in the cells e.g., species KO's and ultimately ability to monitor effects of coupling probabilistic rules to internal marker levels and its effects on growth morphology and growth dynamics. While this was used currently to investigate tumor growth and angiogenesis, a similar set-up with different cell-types could be used to study other multiscale phenomenon.

7.2 Introduction

Breast cancer is one of the prevailing cancers diagnosed among women today as well as the second leading cause of cancer death in women [232]. At the molecular level, signaling networks implicated in breast cancer exhibit considerable complexity in size and connectivity [390, 106, 603, 529, 222]. The intrinsic complexity of these networks complicates the interpretation of experimental findings [603]. Added to this is the dynamical interactions between the cells and their surrounding microenvironmental changes. Extrapolating the molecular signaling events to tissue-level/organ-level phenotypes adds another layer of complexity across spatio-temporal and functional scales [600]. While cellular networks provide deductive capabilities in investigating biological mecha-

nisms, there is a need to inductively extrapolate the consequences of cell signaling to tissue/organ function. This presents an emerging importance and need of models taking both signaling networks and different biochemical and biophysical contexts into account [11, 310, 326, 434, 578]. Computational multiscale modeling approaches are poised to address these growing needs [578]. Multiscale models have been used extensively in several disease states e.g. in cardiac physiology [600], epidermal wound healing [514], renal physiology [112] to identify biological mechanisms, infer multiscale consequences to cell signaling activities.

At the molecular level, human epidermal growth factor receptor 2 (HER2 or ErbB-2) and the associated epidermal growth factor receptor (EGFR) signaling is implicated widely in human breast cancer [585, 624, 213, 335]. Several studies have further implicated a connection between HER2 signaling and vascular endothelial growth factor (VEGF) [421, 416, 317, 72, 446] and interleukin 8 (IL-8) [446, 72] signaling (Fig. 7.1). VEGF signaling plays a role in endothelial cell proliferation, migration as well as alteration of the extracellular environment [130, 125], thereby affecting the survival and invasiveness of breast cancer cells [427, 187]. At the same time, IL-8 signaling promotes growth and metastasis of tumors and angiogenesis [140, 35, 585]. Several studies have reported elevated expression of these pro-angiogenic cytokines in case of aggressive tumor growth and decreased survival of patients with breast cancer [72, 446] highlighting the key role of angiogenesis. Modeling signaling networks involved in breast cancer has come a long way over the years (recently reviewed in Lazzara *et. al.*) [287]. Across a multiscale level, several recent advances have been made in modeling breast cancer [531, 310, 326, 435, 434, 509, 314, 495]. While some studies focused on identifying the role of point-mutations in EGFR signaling in the onset of oncogenic transformations [314, 495], others used this integrative

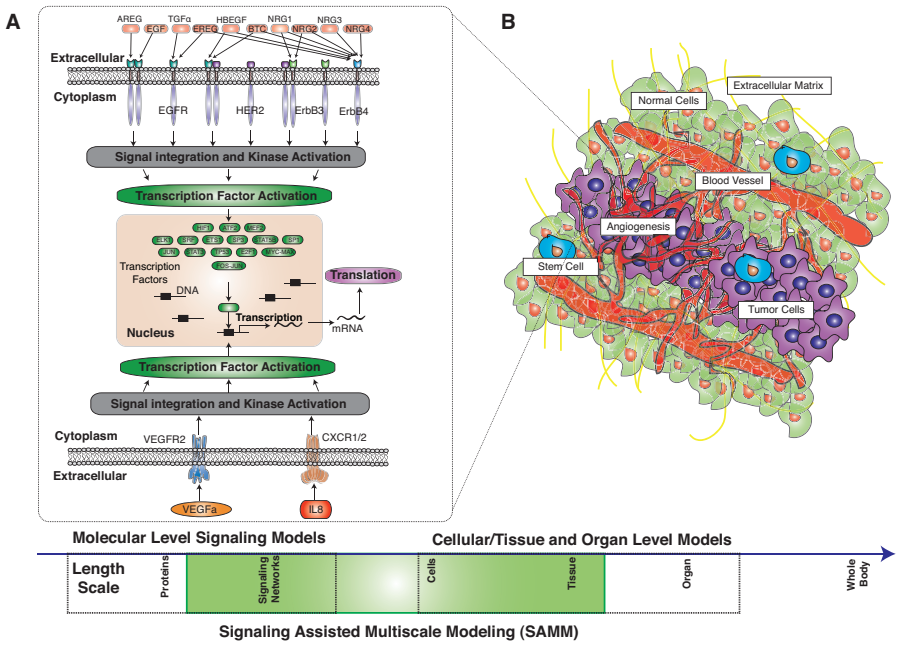


Figure 7.1: (A) Signaling networks implicated in Breast Cancer: EGFR Signaling, VEGF Signaling and IL8 Signaling. These signaling cascades independently or in combination leads to activation of kinases which further downstream leads to activation of transcription factors. These active transcription factors lead to transcriptional regulation of genes in the nucleus. Arrows indicate direction of signal flow. A representative transcription factor is shown to demonstrate transcription. (B) Cancer is a complex process bridging spatial and temporal scales involving many different cell types and processes. The interplay of these cells, extracellular environment and other cues orchestrate cancer growth and progression. Multiscale modeling incorporating multiple processes over wide range of scales in a quantitative manner holds great promise in unraveling the mysteries in case of cancer.

method along with pharmacokinetic measurements to predict resistance based on specific tumor properties and thus improve treatment outcome [143].

We have created a signaling assisted multiscale modeling approach (SAMM) to study systems across multiple length scales. Key highlights of SAMM include use of detailed cellular models to regulate cellular decisions, use of an ensemble

of models to incorporate differences between cells, coupling intracellular markers and probabilistic rules to monitor cellular fate and cellular motility and a hybrid approach to improve speed of solving of external and internal mass-balances. SAMM allows us the flexibility to include multiple cell types, freedom of seeding cells with specific internal cellular machinery within complex geometries, regulate multiple cellular decisions using internal and secreted cellular markers, ease of computation and improved solving speed. In the current study, we demonstrated using SAMM the ability to capture and identify spatio-temporal effects upon growth, ability to model and monitor effects of internal variations in the cells e.g., species KO's and ultimately ability to monitor effects of coupling probabilistic rules to internal marker levels and its effects on growth morphology and growth dynamics. For this study we used a tumor/endothelial cell model which has been used earlier to study eukaryotic hypoxic response [70]. Both VEGF and IL8 have been shown to regulate, in an autocrine signaling based mechanism, the ability of the tumor to grow. While this was used currently to investigate tumor growth and angiogenesis, a similar set-up with different cell-types could be used to study other multiscale phenomenon.

7.3 Modeling Breast Cancer across Multiple Length Scales

Across multiple length scales, several recent advances have been made in modeling breast cancer [531, 310, 435, 434, 509, 314, 495]. While some studies focused on identifying the role of point-mutations in EGFR signaling in the onset of oncogenic transformations [314, 495], others used this integrative method along with pharmacokinetic measurements to predict resistance based on specific tumor properties and thus improve treatment outcome [143]. Efforts in this front

however have faced several challenges ranging from modeling complex intracellular signaling events to spatio-temporal and functional complexities at the tissue-level/organ-levels and crosstalk between the same.

At the molecular level, the number of signaling components involved in the regulation and progression of breast cancer is just immense [222, 197, 350, 209, 106]. However, the key signaling components implicated so far revolve around the role of EGFR, VEGF and IL8 signaling [585, 624, 213, 335, 421, 416, 317, 72, 446, 72]. Malfunctions in EGFR signaling has been highly implicated in several forms of cancer [585, 365, 97, 98, 381, 383, 382]. HER2 (ErbB-2 or c-neu) is the most potent oncoprotein in the family and overexpression of HER2 ($HER2^{+}$) is seen in about 20-30% of breast cancer cases and has been associated with worse prognosis, large tumor size, higher metastatic potential, and decreased survival [476, 624, 213, 335, 616]. Growth and progression of tumors is mediated by the delicate balance of proangiogenic and antiangiogenic factors which regulate angiogenesis [585, 181]. VEGF is one of most potent proangiogenic factors stimulating endothelial cell proliferation and migration and regulating invasiveness of breast cancer cells [125, 427, 187]. A tight relationship between HER2 and VEGF signaling cascades has been seen in preclinical studies on cancer cell lines [421, 416, 317]. It is suggested that these pathways are functionally linked, leading to the hypothesis that VEGF might have a role in the acquired resistance to anti-ErbB drugs when these receptors are pharmacologically blocked [544]. However, regulation of VEGF signaling cascades in breast cancer remain controversial [585]. Another key regulator of angiogenesis is IL-8 [446, 72]. Several studies have reported elevated expression of these pro-angiogenic cytokines in case of aggressive tumor growth and decreased survival of patients with breast cancer [72, 446]. Besides the direct signaling connections, hypoxia in solid tu-

mors (marked by reduced oxygen levels) also stimulates the expression of both IL-8 and VEGF, and VEGF can further enhance IL-8 expression [453]. Modeling signaling networks involved in breast cancer has come a long way over the years (recently reviewed in Lazzara *et. al.*) [287].

At the microscopic scale, coupling internal cellular signaling to dynamical interactions between cells and their surrounding microenvironmental changes and extrapolating this to tissue-level/organ-level phenotypes has seen a lot of activity in the recent years [531, 310, 326, 435, 434, 509, 314, 495]. While some focussed on a holistic picture of cancer modeling several focused on smaller sub issues. Given the central role of EGFR signaling in cancer, Athale *et. al.* studied the effects of EGFR density on tumor growth dynamics, both on the sub- and the multi-cellular levels [17]. They used this to identify links between EGFR density and increase in phenotypic switching activity between proliferative and migratory traits. Using an EGFR gene-protein interaction network module, Zhang *et. al.* investigated the *proliferation-to-migration* switch during tumor growth in case of brain cancer using multiscale approaches [631]. Macklin *et. al.* used an agent based multiscale modeling approach to investigate ductal carcinoma in situ (DCIS) of the breast [328]. They used patient-specific molecular and cellular measurements to calibrate their model as well. Frieboes *et. al.* used multiscale modeling techniques to identify specific functional relationships linking tumor growth and regression to the underlying phenotype under drug response [143]. However, investigating/modeling tumor or any growth shape and morphology is an area full of surprises. In a recent study Engler *et. al.* used multiscale modeling to investigate how emergent properties of adhesion-directed multicellular structures sculpt the tissue, promote its functionality, and maintain its homeostasis through spatial segregation and organization of anchored proteins and

secreted factors and through emergent properties of tissues, including tension fields and energy optimization [119]. Given the innumerable forms of uncertainties and unknowns, Nagl *et. al.* coupled objective bayesian nets to integrate wide sources of information to explore this form of modeling [362]. With respect to angiogenesis, a object-oriented module-based computational integration strategy to build multiscale simulation scheme was recently used by Liu *et. al.* to study exercise induced angiogenesis in skeletal muscles. State of art of modeling multiscale phenomenon has even reached the form of software platforms, e.g., Eissing *et. al.* recently created a computational systems biology software platform for modeling and simulation of multiscale phenomenon [112]. They focused on integrating molecular scale reaction networks to disease biology to whole-body physiology.

7.4 Signaling Assisted Multiscale Modeling (SAMM)

We have created a signaling assisted multiscale modeling technique (SAMM) to study systems across multiple length scales. Key highlights of SAMM include use of detailed cellular models to regulate cellular decisions, use of an ensemble of models to incorporate differences between cells coupling intracellular markers and probabilistic rules to monitor cellular fate and cellular motility and a hybrid approach to improve speed of solving of external and internal mass-balances. SAMM allows us the flexibility to include multiple cell types, freedom of seeding cells with specific internal cellular machinery within complex geometries, regulate multiple cellular decisions using internal and secreted cellular markers, ease of computation and improved solving speed (complete details in materials and methods).

For the current study, at the molecular level, we incorporated a detailed signaling network involving HIF1, VEGF and IL8 signaling cascades into the multiscale framework. A similar strategy has been used successfully before to study non-small cell lung cancer [576, 577, 578] and angiogenesis [326, 310]. The signaling network (describing 548 protein or mRNA species interconnected by 920 interactions) has been used individually before to study the temporal phenomenon involved in VEGF and IL8 signaling in hypoxia [70]. The molecular level of the model was modeled using mass action kinetics within an ordinary differential equation (ODE) framework. The model described the integration between extracellular O_2 levels and VEGFA, IL-8 autocrine induced intracellular gene expression, proliferation and death programs. We modeled the expression and regulation of several transcription factors: HIF1, AP1, ATF2, MEF2 and NF-$\kappa\beta$. AP1 dependent CYCLIND expression was used as a marker of proliferation [557]. NF-$\kappa\beta$ dependent BCL2 expression was used to characterize cell survival while HIF-1α dependent BNIP3 expression was used as a cell death marker [523, 58]. In addition to these markers, the expression of VEGFA, IL-8 and their respective surface receptors VEGFR2 and CXCR1/2 was also modeled. These receptors were connected to transcription factor activation through the activity of the MAPK, p38MAPK and PKC kinases. Complete details of the interactions can be found in Chakrabarti *et. al.* [70] and supplementary materials.

At the microscopic level, a biochemical microenvironment was constructed to represent a virtual tissue sample in either two-dimensional (2D) (data not shown) or three-dimensional (3D) space. In the current study, we used a 20×20×5 grid to demonstrate SAMM. Depending on the need, homogeneous/heterogeneous environments could be attained by distributing external

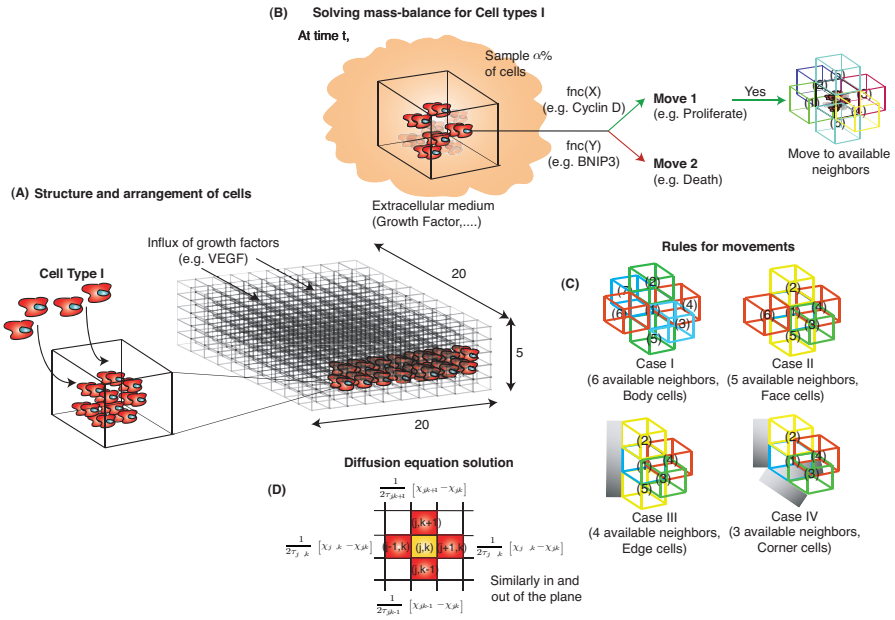


Figure 7.2: Overview of SAMM: (A) A 3D virtual microenvironment comprising of a discrete space with 20×20×5 grid points was created. In this grid, we placed the growth factor source and seeded cells at specific positions at the start of the simulation. In the current study, the GF source was a constant source ($\mathbf{G_{input}}$, in the current study $\mathbf{G_{input}}$ = 10 A.U.) of VEGF and/or IL-8 from the side face as shown. The tumor/endothelial cells were seeded in the opposite face, with a maximum of $\mathbf{N_{Cmax}}$ ($\mathbf{N_{Cmax}}$ = 20 in the current study) per grid element. Each cell in a grid element has a self-maintained signaling network with a randomly assigned parameter vector from the ensemble (details in materials and methods). (B) Over the course of the simulation, cells respond to cellular and environmental cues and accordingly adjust their phenotype at each time step. Intracellular protein marker levels coupled with probabilistic rules were used to characterize proliferation, survival and death phenotypes. AP1 dependent CYCLIND expression was used as a marker of proliferation [557]. NF-$\kappa\beta$ dependent BCL2 expression was used to characterize cell survival while HIF-1α dependent BNIP3 expression was used as a cell death marker [523, 58]. (C) Cellular mobility decisions additionally took into account both number of available neighbors and the cellular densities at neighbors. (D) External diffusive chemical cues (VEGF, IL-8, O_2 in the current study) are regulated via initial seeding, source terms and consumption/production via cells. Throughout the simulation, the three chemical cues are continuously updated at a fixed rate.

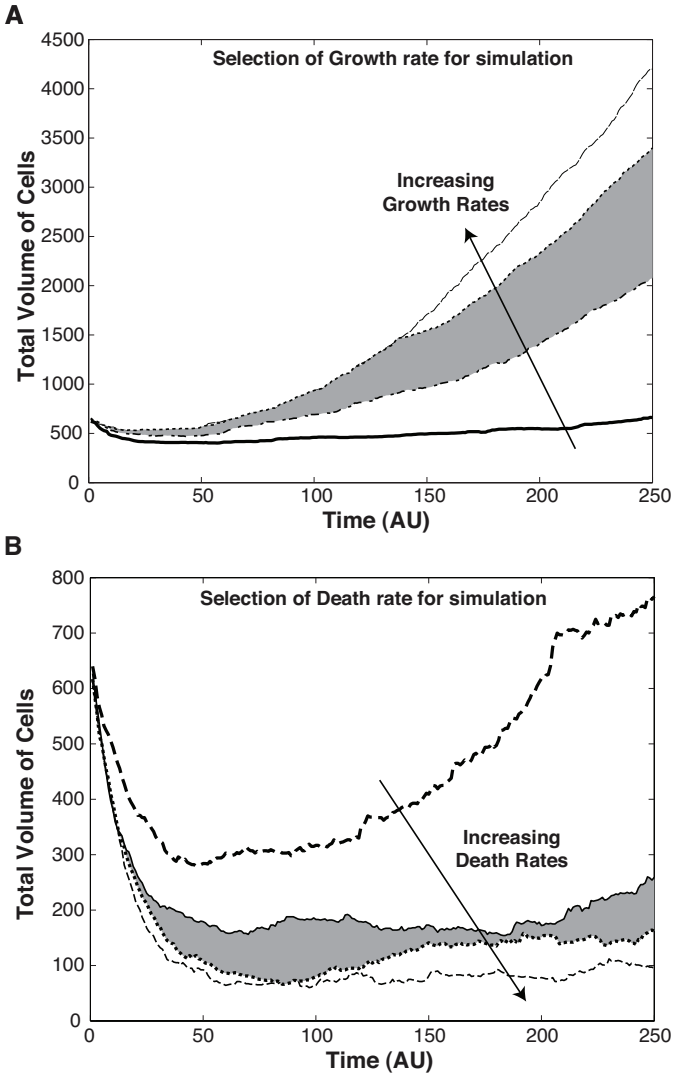


Figure 7.3: Selection for Growth and Death rates: (A) According to the cellular phenotypic "decision" algorithm (discussed in the materials and methods), growth was dependent on the levels of CYCLIND (phosphorylated ERK (ppERK) in one simulation) and a factor G_α. We sampled four orders of magnitude variations in G_α. Final selection of G_α was made so as to have growth behavior in the shaded grey region. (B) Similarly, cell death decision was dependent on the intracellular levels of BNIP3 and a factor D_α. We sampled four orders of magnitude variations in D_α. Final selection of D_α was made so as to have death behavior in the shaded grey region. Both selections were done with the other decision (growth/death) happening as an unbiased random move.

diffusive chemical cues (e.g. growth factors, oxygen) throughout the microenvironmental lattice. We have the freedom/ease in controlling the sources, flow characteristics of these chemical cues. For example, in the current study we had VEGF and/or IL8 flow in at a constant rate ($\mathbf{G_{input}}$ = 10 A.U. details in materials and methods) from one of the side faces (see Fig. 7.2) and we allowed it to diffuse within the above mentioned geometry. Unless otherwise stated, we allowed the oxygen levels to represent normoxic levels (~ 17%) throughout the construct. This allowed us to investigate the isolated VEGF/IL8 induced effects on endothelial/tumor cell behavior. Throughout the simulation, the concentration of the chemical cues were continuously updated at a fixed rate. Generally these diffusion process are implemented based on PDEs [578], however we used a discretized approach to solve for extracellular states (materials and methods), which allowed for reduced simulation time. At any position in the construct, we could have multiple cells of the same kind and/or multiple cell types as well. The number of cells at any position was regulated by the balance of initial seeding, migration from neighbors, growth/death. Each cell type carried a self-determined signaling pathway, as well different set of rate constants which allowed us to capture different phenotypic behaviors as the simulation progressed. Throughout the simulation, the internal states of the cells and resulting cellular counts were continuously updated at a fixed rate. We sampled F_α (0.25 in this study) fraction of the cells at any time (materials and methods). In the current work, molecular and microscopic linkage was based on previously used, experimentally supported molecularly driven cellular phenotype and mobility algorithm [578].

7.4.1 Results

Modeling growth and flow profiles under an external dose of VEGF/IL8:

Using the tumor/endothelial cell model, we investigated growth/death dynamics and migratory behavior of cells upon a dose of growth factors (GF). Constant source of VEGF and/or IL8 was allowed to diffuse from one side face of the construct (grid of 20×20×5) as shown in Fig. 7.2-7.4. The construct was seeded with an initial basal amount of growth factor. One layer of cells were seeded at the opposite face of the construct (maximum of 20 cells per grid element) in the beginning of the simulation (Fig. 7.2 and Fig. 7.4A). During the simulation, cellular fate was decided using internal cellular marker levels: BNIP3 levels for cell death, BCL2 levels for cell survival and CYCLIND levels for proliferation. We narrowed down these rules to select the optimal balance of number of cells growing/dying per unit time by testing a broad range (4 orders of magnitude variations) of probability functions (materials and methods, supplementary Fig. S7.3). Cellular movements were further coupled to probability functions based on cell densities at neighbor locations. Upon start of simulation, cells initially responded to basal GF levels in the construct. This elicits a GF mediated autocrine signal amplification within the cells leading to cellular growth (Fig. 7.4B time step 125 AU). However with progression in time, the GF diffusing in from the other end elicits a second wave of amplified response marked by increased growth (Fig. 7.4B time step 250 AU). Similar behavior with differences in time scale is seen for a dose of VEGF (Fig. 7.4 Fig. 7.8) and IL8 (supplementary Fig. S7.5 column B). As a control, we simulated two cases: a) GF receptor inhibition (Fig. 7.4C) b) No external dose of GF coupled with no initial seeding of GF (Fig. 7.4D). In both cases we see no visible growth of cells (Fig. 7.4C-D).

Subsequently, we used this set-up to study the effects of intracellular variations, variations in external probabilistic rules upon the morphology and dynamics of growth. In the current study, we used only one cell type seeded in a simple grid to demonstrate functionality and utility of SAMM. However a similar approach can very well be used to analyze a wide range of disease states.

Spatio-temporal phenotypes:

As compared to wet experiments, a multiscale computational model allows us the freedom/ease to investigate phenotypic differences between cell populations upon growth (Fig. 7.6). To demonstrate this point, we used the base system (VEGF dose) to analyze the functional states of the cells at three positions (abbreviated as P1, P2 and P3 in Fig. 7.6A) and identify differences if any. We looked at the averaged (average of grid position and its neighbors (max 6)) internal and external marker levels as seen in Fig. 7.6B-F. At P1, the cells in average were exposed to higher levels of VEGF as compared to cells at P3 (Fig. 7.6B), thereby eliciting accordingly a higher VEGF induced IL8 expression at P1 as compared to P3 (Fig. 7.6F). Interestingly, cells at P2 see higher levels of VEGF as compared to cells at P3, yet they elicit a lower VEGF induced IL8 expression (Fig. 7.6F) highlighting the importance of increased time spent by cells at P3 at compared to cells at P2. This suggested that VEGF induced response of the cells was a spatio-temporal phenomenon, which led to non-intuitive behavior of cells at different positions. By comparing the levels of survival (BCL2 Fig. 7.6E), death (BNIP3 Fig. 7.6D) and proliferative (CYCLIND Fig. 7.6C) markers we could see how cells at P2 were the ones having the most action (higher levels as compared to P1 and P3). This suggested that the region P2 was poised at a

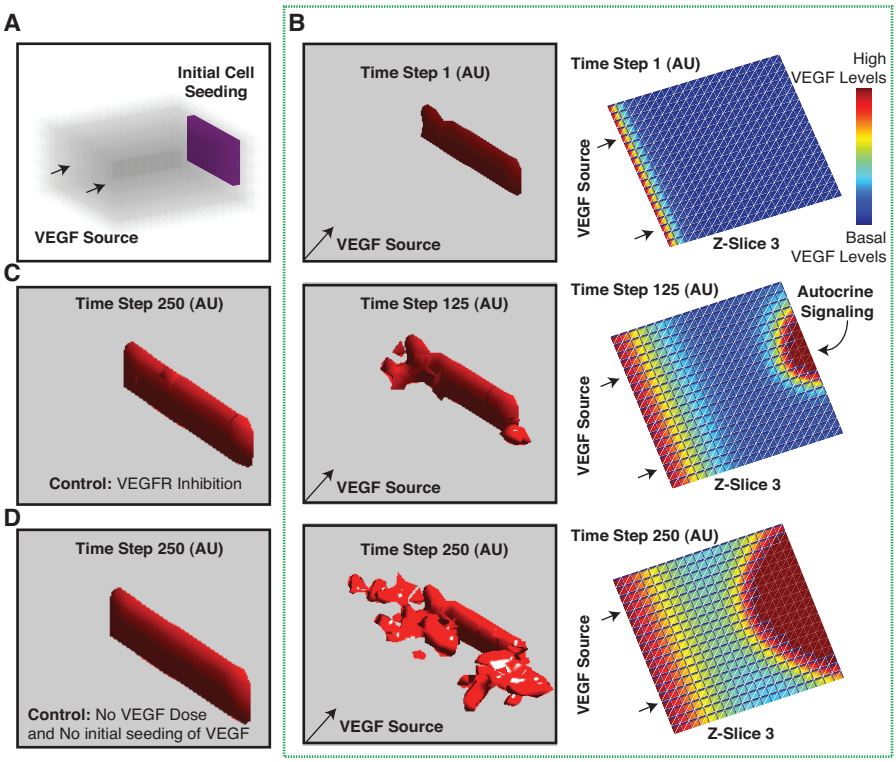


Figure 7.4: Investigating growth/death dynamics and migratory behavior of cells upon a dose of growth factors (GF) using SAMM. (A) Constant source of VEGF and/or IL8 was allowed to diffuse from one side face of the construct (grid of 20×20×5) while tumor/endothelial cells (maximum of 20 cells per grid element) were seeded on the opposite face in the beginning of the simulation. (B) Base case simulation with VEGF (10 au dose, details in materials and methods with rule 2): Growth profile is shown along the left part of panel B, while the external GF (VEGF) levels are shown in the right part of panel B. First row in panel B corresponds to time step 1 while last row in panel B corresponds to time step 250 (end of simulation). (C) Control 1: Simulation with GF (VEGF) receptor inhibition in all cells/at every location in the construct. (D) Control 2: Simulation with no external dose of GF (VEGF) coupled with no initial seeding of GF in the construct.

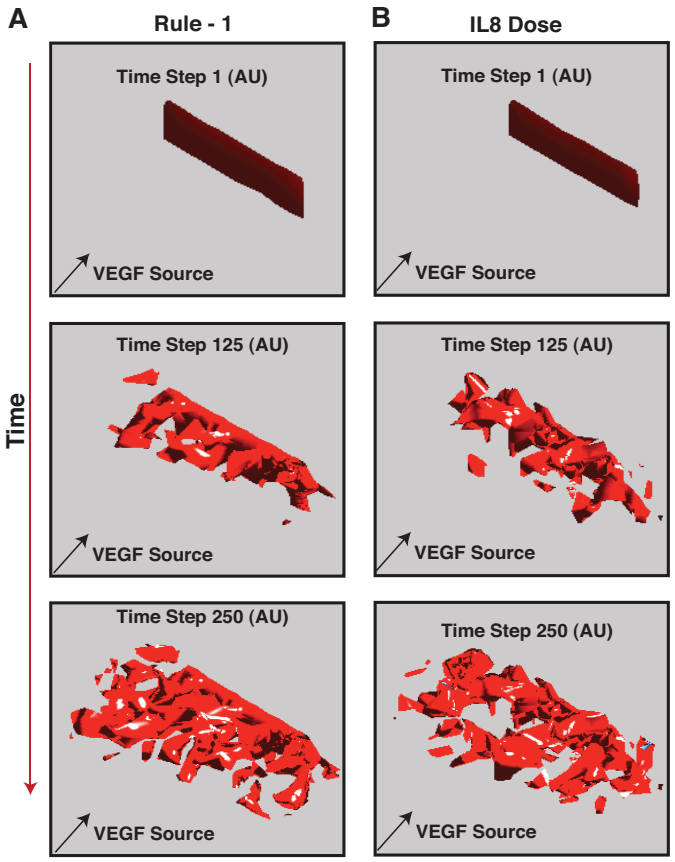


Figure 7.5: Supporting simulations: (A) Growth profile upon VEGF dose with moves based on at least one empty neighbor. (B) Growth profile upon IL8 dose to demonstrate crosstalk between VEGF and IL8 signaling.

delicate balance with optimum VEGF levels and optimum amount of time spent in the given environment thereby acting as a *proliferative core*. This was further substantiated by the additional branches initiating from this region (P2). Interestingly, BNIP3 levels at P1 and P3 were quite similar (Fig. 7.6D) showing that higher levels of VEGF for shorter time (P1) and lower levels of VEGF for longer time (P3) both elicited similar expression of BNIP3. However, BCL2 levels at P1 were higher as compared to P3. This suggested that cells at the tip of a growth were more pro-survival as compared to the cells at the base (P3). It has indeed been shown in the literature that cells at the tip are specialized pro-survival cells,

which utilize dynamic interactions between VEGF and DLL4/Notch signaling cascades at the cellular level to regulate number of tip cells, vessel sprouting and branching [202, 201, 229].

Effects of intracellular perturbations:

Use of detailed molecular networks to guide phenotypic decisions presents us with the scope to analyze the effects of intracellular changes and perturbations to growth. In the base simulations we looked at the growth profile upon a VEGF dose under normoxic conditions (O_2 levels $\sim$ 21%) (Fig. 7.4). However under hypoxic conditions (O_2 levels $\leq$ 2%), cells respond to the reduced oxygen levels via hypoxia inducible factor 1 α (HIF1α) and reactive oxygen species (ROS) [300, 355]. HIF1α signaling leads to the expression of many factors including VEGF and IL-8 [580]. Additionally ROS promotes NF-$\kappa\beta$ activation [355] which further regulates both VEGF and IL-8 expression [355, 580]. Secreted VEGF and IL-8 signals can then amplify the hypoxic response and induce proliferative signals via both autocrine or paracrine signaling mechanisms through the VEGFR and CXCR receptor families. We explored the effects of reduced oxygen levels (O_2 levels $\leq$ 2%) (Fig. 7.7 column A). We saw enhanced growth due to further amplified VEGF and IL-8 signaling mediated by the HIF1 and ROS signaling axis. Autocrine mediated signal amplification is critical in case of VEGF and IL-8 signaling. This positive autocrine feedback at the cellular level is mediated by the activation of AP-1 and NF-$\kappa\beta$ downstream of both VEGF and IL-8 signaling cascades [70]. We simulated a case to verify this by altering the cellular machinery to prevent autocrine signaling (Fig. 7.7 column B). We saw slower and lesser growth overall as compared to the base case (Fig. 7.4 and Fig. 7.7). To suc-

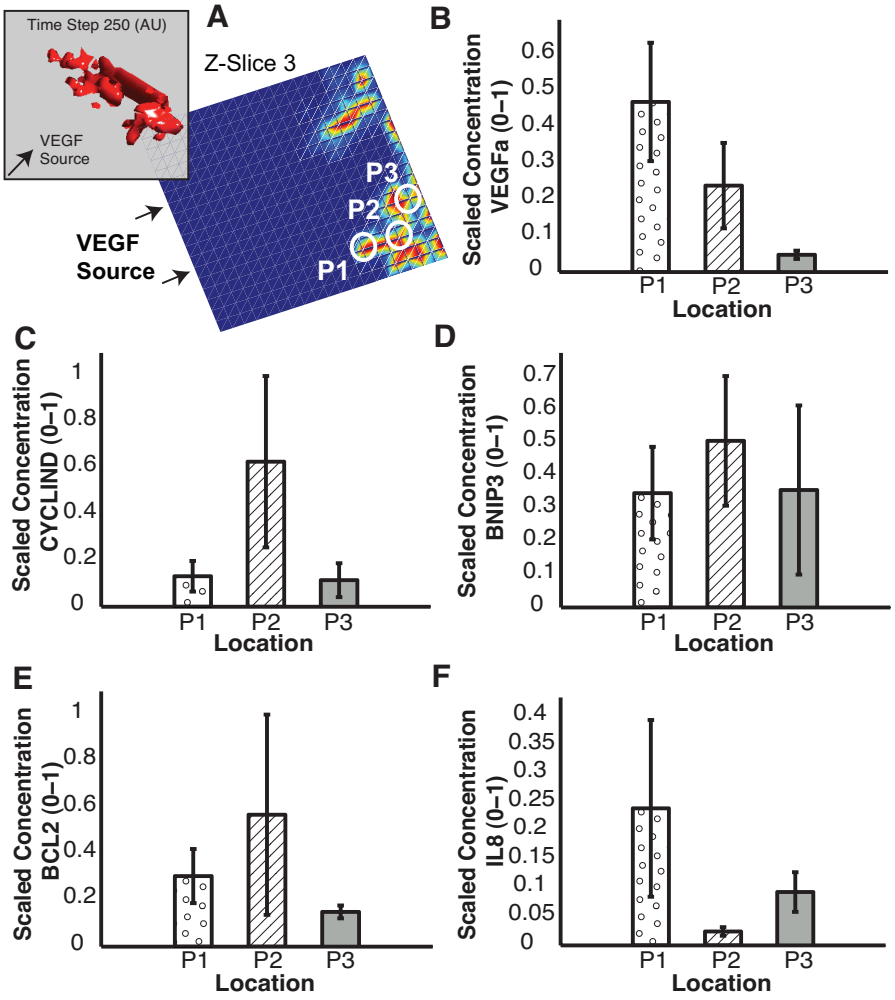


Figure 7.6: Spatio-temporal phenotypic differences during growth: (A) Contour plot (interpolated colors) of cells across the 3^{rd} z-slice at the last time point of simulation (Dose: VEGF, Rule 2). We look at three positions (abbreviated as P1, P2 and P3) to analyze the functional states of the cells. (B)-(F) Averaged (average of the position and its 6 corresponding neighbors) internal and external marker levels at P1, P2 and P3. (B) Scaled extracellular VEGFA levels. (C) Scaled intracellular CYCLIND (marker for proliferation) levels. (D) Scaled intracellular BNIP3 (marker for cell death) levels. (E) Scaled intracellular BCL2 (marker for cell survival) levels. (F) Scaled extracellular IL8 levels.

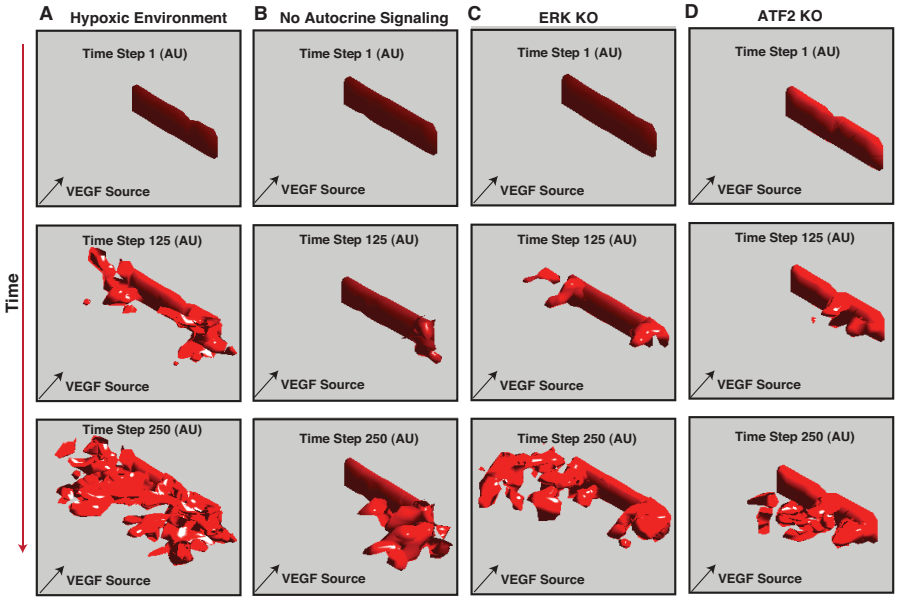


Figure 7.7: Exploring variations in growth profile by altering internal make-up of the cells: Column A - Growth profile upon VEGF dose under hypoxic conditions (O_2 levels $\leq$ 2%). Column B - Growth profile upon VEGF dose with no autocrine signaling within cells. Column C - Growth profile upon VEGF dose with ERK KO in cells. Column D - Growth profile upon VEGF dose with ATF2 KO in cells. Images in the first row correspond to initial time point, while images at third row correspond to last time point of simulation.

cessfully manipulate the VEGF/IL-8 mediated cellular response, it is therefore vital to alter key regulatory cores e.g., MAPK signaling module, AP1 regulatory module. We investigated the effects of these key nodes by simulating knockout experiments. Upon running simulations with cells having ERK KO (materials and methods), we saw significantly slower and lesser growth as compared to the base case (Fig. 7.4 and Fig. 7.7). Similarly, in simulations with cells having ATF2 KO, we saw significantly slower, lesser and disordered growth as compared to the base case (Fig. 7.4 and Fig. 7.7). This could be attributed to the key role of ATF2 mediated regulation of AP1 in maintaining the proliferative ability of the cells via regulation of CYCLIND.

Effect of probabilistic rules:

In case of multiscale phenomenon like tumor growth and angiogenesis, besides spatio-temporal variations and inherent functional variations within cells, there is an additional level of variability introduced by how cells respond to different cellular states and how they decide to move around in the matrix. To investigate this aspect, we simulated several scenarios based on coupling internal cellular marker levels to probabilistic moves incorporating neighboring cellular information in the volume. In the first set of simulations, we coupled intracellular CYCLIND levels (used as a marker for cell proliferation) along with the cell densities of neighbor locations to test different growth profiles. Firstly, successful cellular moves (based on CYCLIND levels) were modeled as random unbiased moves with no effect of neighboring cell densities (Rule 0) (Fig. 7.8 column A). This led to bulk growth as seen in Fig. 7.8 column A. This form of growth, with no neighborhood bias is typical of tumor growth [55]. It is important to note here that there was always a directional bias implying that growth was favored up the gradient of the growth factor. In case of angiogenesis, directionality up the gradient of growth factors is coupled with neighboring cellular information [371]. We investigated this idea by exploring number of sparsely populated neighboring positions (materials and methods). When we allowed moves to occur when there was at least one sparsely populated neighbor (Rule 1) (Fig. S7.5 column A), we saw that we transitioned away from bulk growth to more dendritic growth. We did not see any noticeable bias regarding start of these dendrites from the initial seeded cell surface. However, when we allowed moves to occur when there were at least two (Rule 2) (Fig. 7.8 column B) or three (Rule 3) (Fig. 7.8 column C) sparsely populated neighbors, we saw dendritic growth with a bias towards the edges (Rule 2) and corners (Rule 3).

Behavior under different rules was seen consistently over multiple repeats (10 repeats) and clear differences can be seen between different rules (Fig. 7.9). In all simulations, a completely unbiased move was equally probable. To test the effect of choice of internal cellular marker, we ran a simulation wherein successful cellular moves were based on ERK levels (used earlier by Wang *et. al.* [576, 577]) coupled with at least two sparsely populated neighborhoods (Rule 2 with ERK) (Fig. 7.8 column D). We saw a similar growth morphology as in case of Rule 2, but faster growth.

7.4.2 Discussion

We present a signaling assisted multiscale modeling technique (SAMM) to study systems across multiple length scales. Key highlights of SAMM include use of detailed cellular models to regulate cellular decisions, use of an ensemble of models to incorporate population based behavior of cells, coupling intracellular markers and probabilistic rules to monitor cellular fate and cellular motility and a hybrid approach to improve speed of solving of external and internal mass-balances. SAMM allows us the flexibility to include multiple cell types, freedom of seeding cells with specific internal cellular machinery within complex geometries, regulate multiple cellular decisions using internal and secreted cellular markers coupled with probabilistic functions and improved solving speed. This platform allows us to investigate complex multiscale phenomenon e.g., tumor growth, angiogenesis. In the current study, we demonstrated using SAMM the ability to capture and identify spatio-temporal effects upon growth (Fig. 7.6), ability to model and monitor effects of internal variations in the cells e.g., species KO's (Fig. 7.7) and ultimately ability to monitor effects of coupling

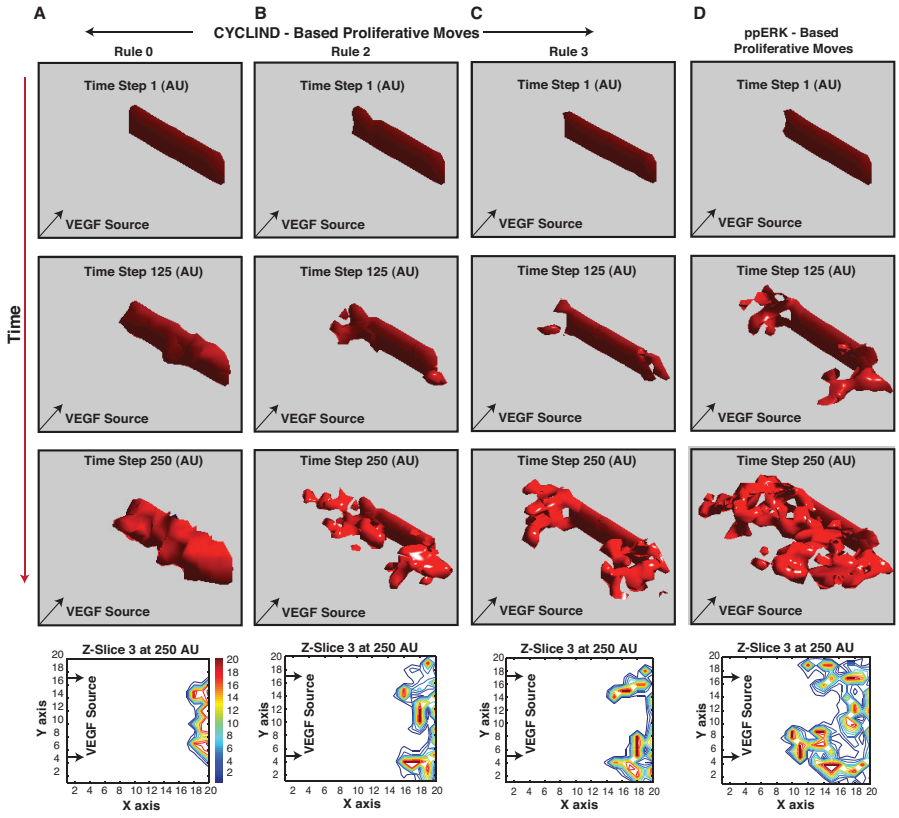


Figure 7.8: Exploring variations in growth profile due to probabilistic rules and alternate use of cellular markers: Column A - Rule 0: Growth profile upon VEGF dose with complete unbiased moves. Column B - Rule 2: Growth profile upon VEGF dose with moves based on at least two empty neighbors. Column C - Rule 3: Growth profile upon VEGF dose with moves based on at least three empty neighbors. Column D - Rule 2 + ERK: Growth profile upon VEGF dose with moves based on phosphorylated ERK levels instead of CYCLIND levels. Images in the first row correspond to initial time point, while images at third row correspond to last time point of simulation. Images in the fourth row correspond to the contour plot of cells across the 3^{rd} z-slice at the last time point of simulation.

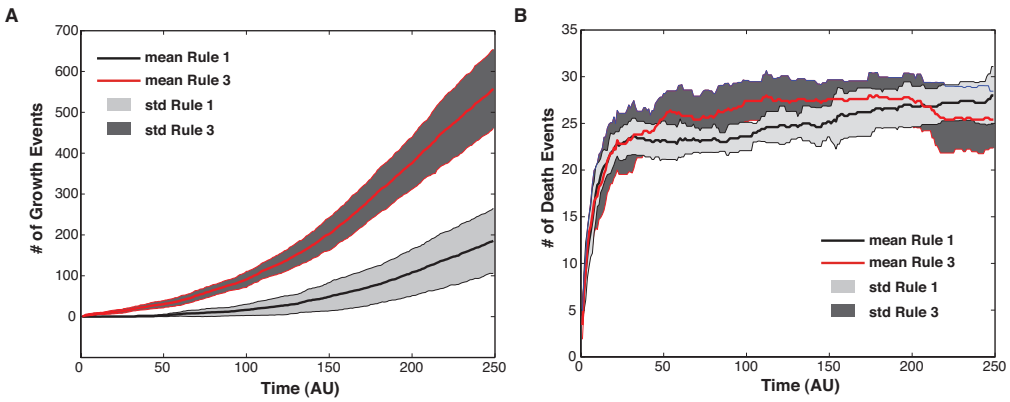


Figure 7.9: Repeatability check: We simulated multiple repeats (10 repeats) for rule 1 and rule 3 to check for reproducibility and consistency. (A) Number of growth events for rule 1 (mean of 10 simulations as black line and shaded light grey area demarcating one standard deviation) and rule 3 (mean of 10 simulations as red line and shaded dark grey area demarcating one standard deviation) over the simulation time. We clearly see differences between the two different rules over multiple repeats while seeing consistent behavior over repeats. (B) Number of death events.

probabilistic rules to internal marker levels and its effects on growth morphology and growth dynamics (Fig. 7.8).

Different regions of a growing mass of cells are bound to have different functionalities and thereby different phenotypic characteristics [19]. For example, in case of angiogenesis tips of the sprouts are actually occupied by endothelial tip cells which dynamically regulate number of tip cells, vessel sprouting and branching [202, 201, 229]. Similar phenomenon is seen even in case of epithelial tubulogensesis [556] and axon guidance [67]. As compared to wet experiments, a multiscale computational model e.g. SAMM allows us the freedom/ease to investigate phenotypic differences between cell populations at different regions of growth (Fig. 7.6). We do not see any linear relationships between cellular states as we move from the tip of a dendrite to the base, suggesting a complex

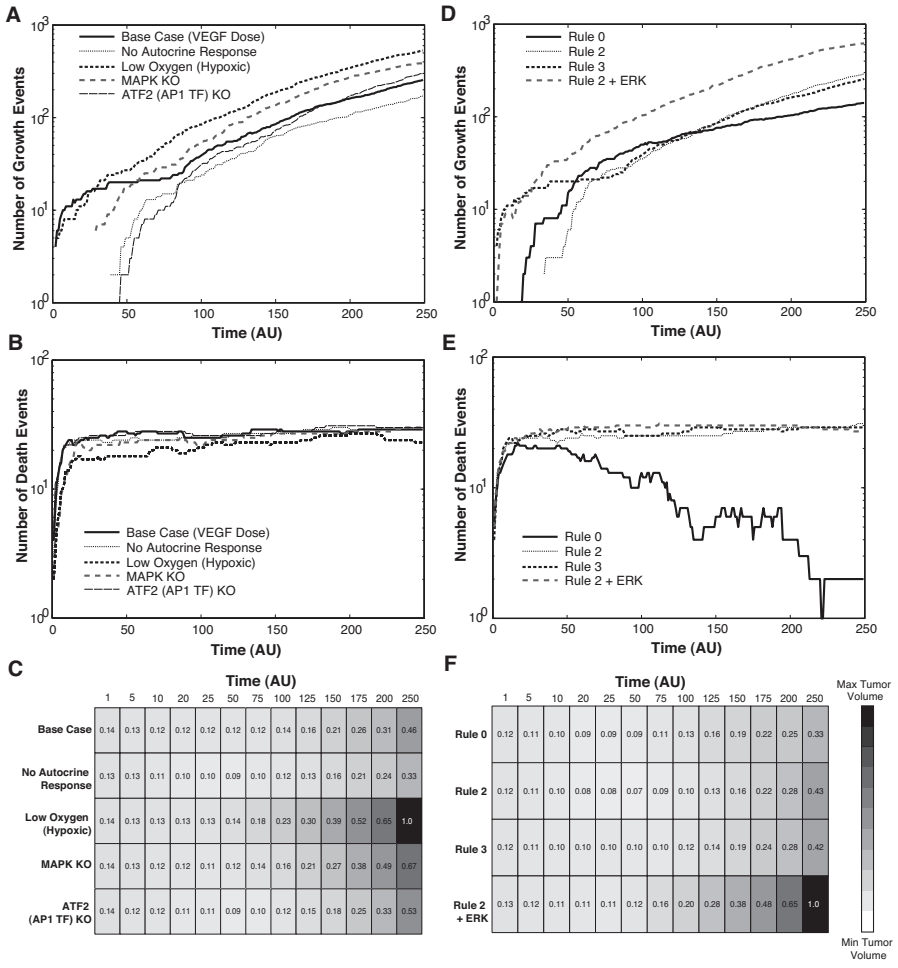


Figure 7.10: Growth and death dynamics across simulations: (A-C) Growth/death characteristics upon intracellular perturbations. (D-F) Growth/death characteristics upon probabilistic rule variations. In general, growth dynamics (linear growth rate) was observed overall consistently throughout each simulation. We do see differences between individual cases with respect to time of start of growth, total volume of growth, e.g., at low O_2 growth was faster and more pronounced due to additional effects mediated by the ROS and HIF1 signaling axis.

spatio-temporal phenomenon at play. Cells at the tip of the dendrite are exposed to higher GF levels and thereby elicit more pronounced GF mediated effects. However there is an added component of how long these cells have actually stayed in this environment. Our simulations suggest that the region towards the center of the dendrite (Fig. 7.6A) is poised at a delicate balance with optimum VEGF levels and optimum amount of time spent in the given environment thereby acting as a *proliferative core*. Also cells at the tip of the dendrite are more pro-survival as compared to the cells at the base. This idea has indeed been recently suggested by Franco *et. al.*, where in they speculate that growing endothelial cells recruit pericytes which ultimately lead to increase anti-apoptotic protein levels to promote survival of cells at the tip [139]. However for accurate predictions and analysis, we do need more detailed/comprehensive intracellular cellular signaling models and additional cell type models e.g., tip cell, stalk cell models. Using different cell type models with inherent differences in machinery might provide a clearer picture regarding heterogeneity and elucidate spatio-temporal variations along different segments of growth.

Regulation of complex multicellular processes like tumor growth, angiogenesis is often done via altering internal cellular makeup directly or indirectly by using external agents e.g., anti-angiogenic drugs. SAMM presents the scope to analyze the effects of intracellular changes/perturbations to growth morphologies/dynamics. Crosstalk between several signaling axis e.g., HIF1 signaling, ROS signaling, VEGF signaling and IL-8 signaling all lead to and control outcome of VEGF/IL-8 mediated response of cells [300, 355, 580]. In this regard, autocrine mediated signal amplification is key to the time/magnitude of the response. We tested several cases by directly/indirectly altering internal states of the cells. While the growth dynamics (linear growth rate) was overall consis-

tent throughout each case (Fig. 7.10A), we did see differences between individual cases (Fig. 7.10A-C). For example, at low O_2 growth was faster and more pronounced due to additional effects mediated by the ROS and HIF1 signaling axis. This behavior is consistent with literature evidence [188, 284, 417, 212, 53]. Similarly, ERK KO and ATF2 (transcription factor for AP1) KO lead to delayed and lower growth as compared to the base case. It does not completely inhibit growth due to other redundant routes of CYCLIND regulation. This demonstrates an example of two strategies which delay growth but do not alter the rate of growth in the long run. In the current scheme however, we did not take several signaling components into account. For example, we did not include any signaling modules associated with cellular metabolism, EGFR signaling and no nutrient dependency was taken into account. These signaling components amongst many others play a critical role in breast cancer progression [588, 251, 74, 222] and angiogenesis [109, 559, 420]. Incorporating intracellular details shall be vital to exploring perturbation based experiments. However, added size and complexity of elaborate signaling models bring with themselves further challenges associated with solution techniques and computing time limitations. Even with the obstacles we are about to face, this form of analysis could be a valuable tool to test the effectiveness of mechanistic regulators of cellular machinery in multiscale phenomenon like tumor growth and angiogenesis.

An interesting area of cellular growth and migration are the resultant morphologies. There could be a diverse range of factors influencing the same e.g., concentrations of ECM proteins, gradients of growth factors, neighboring cellular densities, nutrient levels, physical stresses. SAMM allows us the flexibility to study the influence of these external factors on growth morphologies. Just by incorporating neighboring cellular density information we could transition

from a bulk growth (as seen in tumors [55]) to well formed dendritic growth with/without a bias towards corners and edges (as seen in 3D cultures [371]). In reality there must be more at play. Looking forward, using SAMM we can reverse engineer a particular growth morphology by sampling the probability function landscape. This shall allow us to test the contribution and implications of different extracellular factors in determining the nature of growth. Another area which we barely touched upon in the current study was use of intracellular markers to determine cellular phenotype. For example, we saw differences in growth when we based the proliferative moves based on intracellular levels of CYCLIND and ppERK. This begs further investigation with regards to choice of individual or combinatorial use of intracellular marker levels to model phenotypic decisions. Using SAMM, we can identify the relevance/validity of these choices.

7.5 Conclusion

Unraveling the mysteries within complex multiscale phenomenon like tumor growth and angiogenesis, with its inherent crosstalk between molecular and microscopic scales makes this area a challenging field of study, both experimentally and computationally. Utilizing the current progress in the field of modeling signaling networks, modeling complicated 3D constructs at tissue/organ/body levels we are at the perfect position to make strides of progress in creating multiscale models incorporating information across multiple length scales. Modeling such complicated phenomenon has come a long way over the years [485, 110, 500, 418, 154]. Across a multiscale level, several recent advances have been made in modeling breast cancer [531, 310, 326, 435, 434, 509, 314, 495, 17,

631, 328, 143, 119, 362, 112]. While some studies focused on identifying the role of point-mutations in EGFR signaling in the onset of oncogenic transformations [314, 495], others used this integrative method along with pharmacokinetic measurements to predict resistance based on specific tumor properties and thus improve treatment outcome [143]. We present a signaling assisted multiscale modeling technique (SAMM) to study systems across multiple length scales. Key highlights of SAMM include use of detailed cellular models to regulate cellular decisions, use of an ensemble of models to incorporate population based behavior of cells, coupling intracellular markers and probabilistic rules to monitor cellular fate and cellular motility and a hybrid approach to improve speed of solving of external and internal mass-balances. SAMM allows us the flexibility to include multiple cell types, freedom of seeding cells with specific internal cellular machinery within complex geometries, regulate multiple cellular decisions using internal and secreted cellular markers coupled with probabilistic functions and improved solving speed. This form of a platform allows us to investigate complex multiscale phenomenon e.g., tumor growth, angiogenesis, relevant to several complex disease states like breast cancer.

7.6 Materials and Methods

7.6.1 Molecular scale: Formulation and solution of model equations for internal cellular signaling.

The tumor/endothelial cell model was adapted from a previous study by Chakrabarti *et. al.* and described 548 protein or mRNA species interconnected

by 920 interactions [70]. The model described the integration between extracellular oxygen (O_2) levels and VEGFA (vascular endothelial growth factor A), IL-8 (interleukin 8) autocrine induced intracellular gene expression, proliferation and death programs. For HIF-1α (hypoxia inducible factor 1 α), we modeled oxygen-dependent PHD (prolyl-4-hydroxylase)/VHL (von hippel-lindau)-induced HIF-1α degradation and FIH (factor inhibiting HIF)-induced inhibition. In addition to HIF-1α, we modeled the expression and regulation of four other transcription factors: AP1 (activator protein 1), ATF2 (activating transcription factor 2), MEF2 (myocyte enhancer factor-2) and NF-$\kappa\beta$ (nuclear factor $\kappa\beta$). AP1 expression was dependent upon both ATF2 and MEF2. Both ATF2 and MEF2 were constitutively expressed and regulated by phosphorylation at T69/T71/T73 for ATF2 and T312/T319 for MEF2 by p38MAPK family members. We modeled a single activating phosphorylation site on AP1 (S63/S73), which was regulated by ppERK (phosphorylated ERK) which is formed downstream of both IL-8 and VEGFA signaling cascades. NF-$\kappa\beta$ regulation was more complex; NF-$\kappa\beta$ was assumed to be sequestered by a lumped pool of IκB family members. The regulation of this pool occurred by phosphorylation at S32/S36 and Y42 on IκB; S32/S36 was phosphorylated by PKC (protein kinase C) while Y42 was regulated by an unknown ROS-dependent kinase. Lastly, protein markers downstream of these transcription factors were used to characterize proliferation, survival and death phenotypes. AP1 dependent CYCLIND expression was used as a marker of proliferation. NF-$\kappa\beta$ dependent BCL2 expression was used to characterize cell survival while HIF-1α dependent BNIP3 expression was used as a cell death marker. In addition to these markers, the expression of VEGFA, IL-8 and their respective surface receptors VEGFR2 and CXCR1/2 was also modeled. These receptors were connected to transcription

factor activation through the activity of the MAPK, p38MAPK and PKC kinases. Complete details of the interactions in the model are enlisted in Chakrabarti *et. al.* [70].

The cellular signaling model was formulated as a set of coupled ordinary differential equations (ODEs):

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{r}(\mathbf{x}, \mathbf{p}) \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{7.1}$$

The symbol $\mathbf{S}$ denotes the stoichiometric matrix (548×920). The quantity $\mathbf{x}$ denotes the concentration vector of proteins or protein complexes (548×1). The term $\mathbf{r}(\mathbf{x}, \mathbf{p})$ denotes the vector of reaction rates (920×1). Each row in $\mathbf{S}$ described a protein or protein-protein complex, while each column described the stoichiometry of network interactions. Thus, the (i, j) element of $\mathbf{S}$, denoted by σ_{ij}, described how protein i was involved in rate j. If $\sigma_{ij} < 0$, then protein i was consumed in r_j. Conversely, if $\sigma_{ij} > 0$, protein i was produced by r_j. Lastly, if $\sigma_{ij} = 0$, there was no protein i in rate j. All of these interactions were obtained from the literature.

We assumed mass-action kinetics for each interaction in the network. The rate expression for interaction q was given by:

$$r_q\left(\mathbf{x}, k_q\right) = k_q \prod_{j \in \{\mathbf{R}_q\}} x_j^{-\sigma_{jq}} \tag{7.2}$$

The set $\left\{\mathbf{R}_q\right\}$ denotes reactants for reaction q while σ_{jq} denotes the stoichiometric coefficient (element of the matrix $\mathbf{S}$) governing species j in reaction q. All reversible interactions were split into two irreversible steps. The mass-action formulation, while expanding the dimension of the model, regularized the mathematical structure; this allowed automatic generation of the model code using UNIVERSAL and regularized the unknown model parameters (pa-

rameters were one of only three types, association, dissociation or catalytic rate constants). UNIVERSAL, an open source Java code generator, generates multiple code types from text and SBML inputs. UNIVERSAL is freely available as a Google Code project (http://code.google.com/p/universal-code-generator/). Thus, although mass-action kinetics increased the number of parameters and species, they reduced the complexity of model analysis. In this study, we considered well-mixed nuclear, cytosolic and extracellular compartments. The model equations were solved using the LSODE routine in OCTAVE (v 3.1.0; www.octave.org) on an Apple workstation (Apple, Cupertino, CA; OS X v10.6.4).

Model parameters were estimated using 20 experimental data sets taken from literature (further details in Chakrabarti *et. al.* [70]). In total, 1468 unknown parameters (920 kinetic parameters and 548 initial conditions) were present in the cellular model. Both kinetic parameters and unspecified initial conditions were estimated using the POETs multiobjective optimization algorithm [504]. Pareto Optimal Ensemble Techniques (POETs) is a multiobjective optimization strategy which integrates several local search strategies e.g., Simulated Annealing (SA) or Pattern Search (PS) with a Pareto-rank-based fitness assignment [504]. POETs generated more than 40,000 probable models. For SAMM, we selected models with Pareto rank one or less (N = 3233) similar to the ones used in the previous study [70]. Each cell in the simulation was randomly assigned a parameter set form this ensemble. This ensured diversity between cells in the population and allowed us to explore population average behavior. SBML file of the model along with the ensemble of parameters used for SAMM are included in the supplementary materials.

7.6.2 Microscopic scale: Formulation and evolution of extracellular microenvironment.

As shown in Fig. 7.2, a 3D virtual microenvironment comprising of a discrete space with 20×20×5 grid points was created. In this grid, we placed the growth factor source and seeded cells at specific positions at the start of the simulation. In the current study, the GF source was a constant source ($\mathbf{G_{input}}$, in the current study $\mathbf{G_{input}}$ = 10 A.U.) of VEGF and/or IL-8 from the side face as shown in Fig. 7.2. The tumor/endothelial cells were seeded in the opposite face, with a maximum of $\mathbf{N_{Cmax}}$ ($\mathbf{N_{Cmax}}$ = 20 in the current study) per grid element (Fig. 7.2). Each cell in a grid element has a self-maintained signaling network with a randomly assigned parameter vector from the ensemble as discussed earlier. While in the current study, we demonstrate the functionality using a simple grid based setup, SAMM allows us the flexibility to easily create complex geometries with intricate arrangements of GF sources and seeded cells.

Over the course of the simulation, cells respond to cellular and environmental cues and accordingly adjust their phenotype at each time step. However, at any time step we sampled only F_α (in the current study F_α = 0.25) fraction of cells. Sampling of cells is done randomly, so overall we should capture the trends over the entire area. The fractional sampling is used in order to minimize overall simulation time. A simulation run is terminated after 250 time steps (this corresponds to ~ 10 experimental days). We use the number of live cells at any time for total growth volume (as shown in Fig. 7.10). We also record the number of growth and death events at every grid element, at every time step for further analysis (as shown in Fig. 7.10).

External diffusive chemical cues (VEGF, IL-8, O_2 in the current study) are regulated via initial seeding, source terms and consumption/production via cells. Unless otherwise mentioned, we maintained O_2 levels to represent normoxic conditions to isolate and study the GF mediated response. Throughout the simulation, the three chemical cues are continuously updated at a fixed rate. The concentration of extracellular variables ($\mathbf{C}$) at spatial position (i, j, k) was governed by ordinary differential equations of the form:

$$\frac{d\mathbf{C}_{ijk}}{dt} = \mathbf{D}_C \dot{\nabla}^2 \mathbf{C}_{ijk} + \theta(C), \qquad t = 1, \ldots, T_{end} \tag{7.3}$$

The first-term of Eqn. (7.3) described the diffusive component with $\mathbf{D}_C$ being be corresponding diffusion constant for the extracellular species ($\mathbf{C}$), while the second-term described consumption/production by cells. T_{end} in the current study is 250 which corresponding to $\sim$ 10 experimental days.

We can rewrite this equation as:

$$\frac{d\mathbf{C}_{ijk}}{dt} = \sum_{f=1}^{6} \mathcal{T}_f \left(\mathbf{C}_f - \mathbf{C}_{ijk}\right) + \mathcal{H}\left(\mathbf{C}^* - \mathbf{C}_{ijk}\right) + \theta(C), \qquad t = 1, \ldots, T_{end} \tag{7.4}$$

where $i = 1, \ldots, N_x; j = 1, \ldots, N_y; k = 1, \ldots, N_z$. The first-term of Eqn. (7.4) described the material transport between subvolumes while the second-term described transport of material to and from the boundary and the third term described consumption/production by cells. The advantages of this solution strategy is drastic reduction of simulation time and ease of computation. After each simulation step, we matched internal and external protein marker levels for consistency.

7.6.3 Molecular-microenvironment crosstalk: Molecularly driven cellular phenotype and mobility calculations.

Intracellular protein marker levels coupled with probabilistic rules were used to characterize proliferation, survival and death phenotypes. AP1 dependent CYCLIND expression was used as a marker of proliferation [557]. NF-$\kappa\beta$ dependent BCL2 expression was used to characterize cell survival while HIF-1α dependent BNIP3 expression was used as a cell death marker [523, 58]. Probability of growth $\mathbf{P}_g$ for any cell at time t was modeled as:

$$\mathbf{P}_g = \frac{\mathcal{X}_g}{G_\alpha + \mathcal{X}_g} \tag{7.5}$$

where $\mathcal{X}_g$ refers to intracellular growth marker levels and G_α refers to a growth constant. In the current study, growth was dependent on the levels of CyclinD (ERK in one simulation, similar to Wang *et. al.* study [576, 577]). We sampled four orders of magnitude variations in G_α. Final selection of G_α was made so as to have growth behavior in the shaded grey region as seen in supplementary Fig. S7.3. Selections was done with the death decision happening as an unbiased random move. Higher G_α would mean lower probability of growth and vice-versa. Similarly, probability of death $\mathbf{P}_d$ for any cell at time t was modeled as:

$$\mathbf{P}_d = \frac{\mathcal{X}_d}{D_\alpha + \mathcal{X}_d} \tag{7.6}$$

where $\mathcal{X}_d$ refers to intracellular death marker levels and D_α refers to a death constant. In the current study, death was dependent on the levels of BNIP3. We sampled four orders of magnitude variations in D_α. Final selection of D_α was made so as to have death behavior in the shaded grey region as seen in supplementary Fig. S7.3. Selections was done with the growth decision happening as an unbiased random move.

Furthermore, cellular mobility decisions additionally took into account both number of available neighbors and the cellular densities at neighbors. For example, grid positions were characterized as body, face, edge or corner positions which accordingly determined which available neighbors were available for a move (as illustrated in Fig. 7.2). Cellular densities at neighborhood **i** were calculated as:

$$\beta_{\mathbf{i}} = \frac{\mathbf{N_{Ci}}}{\mathbf{N_{Cmax}}} \tag{7.7}$$

where $\mathbf{N_{Ci}}$ is the number of cells at neighborhood **i**, $\mathbf{N_{Cmax}} = 20$ and **i** ranged from 1 – 6 depending on the grid position. A neighbor was considered to be sparsely populated if $\beta_{\mathbf{i}} < \beta_{\mathbf{T}}$, where $\beta_{\mathbf{T}}$ was the threshold cellular density (in current study $\beta_{\mathbf{T}} \leq 0.1$). As shown in Fig. 7.8 and Fig. S7.5, we simulated several scenarios based on number of sparsely populated neighbors. *Rule 0* corresponded to the case where cellular moves were random unbiased moves with no effect of neighboring cell densities. *Rule 1* corresponded to the case where cellular moves were allowed to occur when there was at least one sparsely populated neighbor. *Rule 2* corresponded to the case where cellular moves were allowed to occur when there were at least two sparsely populated neighbors. *Rule 3* corresponded to the case where cellular moves were allowed to occur when there were at least three sparsely populated neighbors. At all times, we did have the option for a random growth event as well.

7.6.4 Intracellular perturbation analysis and control simulations.

In this study, simulated KO experiments (e.g., ERK KO, ATF2 KO) were carried out by setting corresponding gene levels to zero in the intracellular model for all the cells within the construct. This allowed us to mimic the effects of the an experimental KO in out simulations.

As control simulations, we simulated several cases to isolate and test the validity of the GF mediated response. Sample control experiments included setting diffusion constants $\mathbf{D}_C$ as zero, GF KO, GF receptor inhibition in which we prevented GF receptors from sensing the external GF levels by setting the corresponding rate constants to zero over the entire ensemble, No external dose of GF with/without initial seeding of GF across the construct. Some of the control experiments are discussed in the manuscript (Fig. 7.4C-D).

CHAPTER 8

TOP-DOWN ANALYSIS OF HYPOXIA INDUCED SIGNALING DIFFERENCES IN 2D VS 3D

8.1 Abstract

Recent advancements in high-throughput methods for characterizing genomic, transcriptomic, proteomic, and metabolomic states allow one to view the global consequences of perturbations rather than just the *usual suspects*. High-throughput experimental techniques of gathering disease state information have created the need for specialized modeling strategies to build predictive disease models. In case of hypoxia induced tumor growth and angiogenesis, we employed the top-down strategy to explore the cellular behavior and identify different operational modes of cells. We created an *in-house* platform to analyze microarray data and overlay the quantitative information on pre-curated pathways to identify malfunctions within the cells. We coupled this to mathematical analysis techniques like singular value decomposition (SVD) to identify key mediators of the response. Using these strategies, we identified differential regulators of the cellular behavior in 2D versus 3D experimental protocols. We also identified differential behaviors and involvement of key signaling pathways in mediating the hypoxic response. While this form of analysis provided key insights in the experimental directions, which are being pursued by our collaborators, we specifically used this insight to create proof-of-concept physiochemical models to further the process of modeling these disease states. This strategy has immense scope of development, from identifying key regulators to forming proof-of-concept physiochemical models.

8.2 Introduction

Recent advancements in high-throughput methods for characterizing genomic, transcriptomic, proteomic, and metabolomic states allow one to view the global consequences of molecular perturbations rather than just the *usual suspects* [508]. However, this wealth of *high-throughput* data creates new challenges in data interpretation, because most of the measurements lack a biological context for interpreting the biological relevance to the experimental perturbation. These measurements give a snapshot of activity at various levels at the time of isolation [508, 606, 166]. While this has successfully provided a detailed view in several disease states, e.g. B cell lymphoma [5], breast tumors [426, 419]; it lacks providing us with the understanding of relationships within gene networks [606, 166].

Interactions between gene products are extremely complex and are not the only determinant of cellular behavior [508]. There are positive and negative feedback controls via transcription factors to begin with. To address the area of inferring connectivity from such high-throughput data, several groups are working to set up tools for the trade [136, 52, 51, 96, 95, 503, 594, 528, 169]. Researchers use clustering or correlation based techniques to infer information about the topology of the connections extensively [528, 169, 303, 594, 24, 293]. These techniques help reduce the complexity of these data sets by identifying clusters of signaling species that may either be co-regulated or that can similarly regulate other species in a signaling network. However these techniques are insufficient to determine the kinetic or dynamic parameters required to model the dynamical behavior of the system [508]. Another strategy in extracting potential gene relationships from such high-throughput data is by using singular value decomposition (SVD) [568, 252, 87, 632, 607, 530, 8, 9, 104]. These approaches

draw on the use of computational tools to identify features in the data that may globally represent the entire data set (e.g., principal components). These approaches inherently make few/no assumptions about the data and have the capability to provide unbiased identification of unexpected correlations. These correlations might not explain the causality between correlated components, however they can be useful for identifying nonintuitive patterns in the data and help guide future experiments.

In the current study, using a top-down strategy, we used these extensive data-sets to investigate cellular signaling, identify malfunctions, create predictive models to infer patient outcome. In case of hypoxia induced tumor growth and angiogenesis, we employed this strategy to explore the cellular behavior. We created an in house platform to analyze microarray data and overlay the quantitative information on pre-curated pathways to identify malfunctions within the cells. We coupled this to mathematical analysis techniques like singular value decomposition (SVD) to identify key mediators of the response. Using these strategies, we identified differential regulators of the cellular behavior in 2D versus 3D experimental protocols. We also identified differential behaviors and involvement of key signaling pathways in mediating the hypoxic response. While this form of analysis provided key insights in the experimental directions, which are being pursued by our collaborators, we specifically used this insight to create proof-of-concept physiochemical models to further the process of modeling these disease states. This strategy has immense scope of development, from identifying key regulators to forming proof-of-concept physiochemical models.

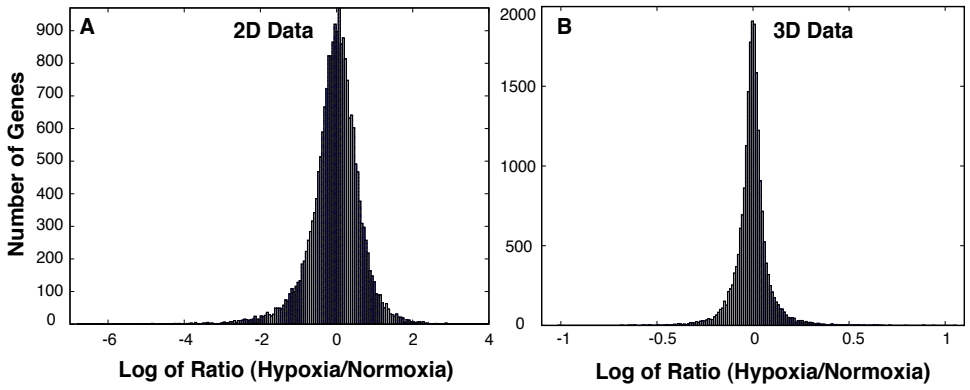


Figure 8.1: Histogram of 2D and 3D fold changes in mRNA components (hypoxia/normoxia): We see a wider distribution in case of 2D experiment as compared to 3D experiment. Details of the experimental protocols are in the materials and methods.

8.3 Results

8.3.1 Big Picture Analysis of the Microarray Data

Oral squamous cell carcinoma-3 (OSCC3) cells cultured in 2D and 3D (material and methods) were exposed to normoxia (17%) and hypoxia (1%) oxygen for 6 days and then microarray measurements were done using Affymetrix GeneChips. Microarray measurements gave us measurements of dimensions (54675×3 (three repeats)) for each experimental condition (hypoxia, normoxia - 2D and 3D). For further analysis, we used the mean of the measurements over the three repeats. As in Fig. 8.1, we saw a wider distribution of fold change (hypoxia versus normoxia) in expression levels in case of 2D experiment as compared to 3D.

Microarray arrays have multiple probes for the same IDs. So the first step of further analysis was to remove repetitions in order to reduce dimensions of

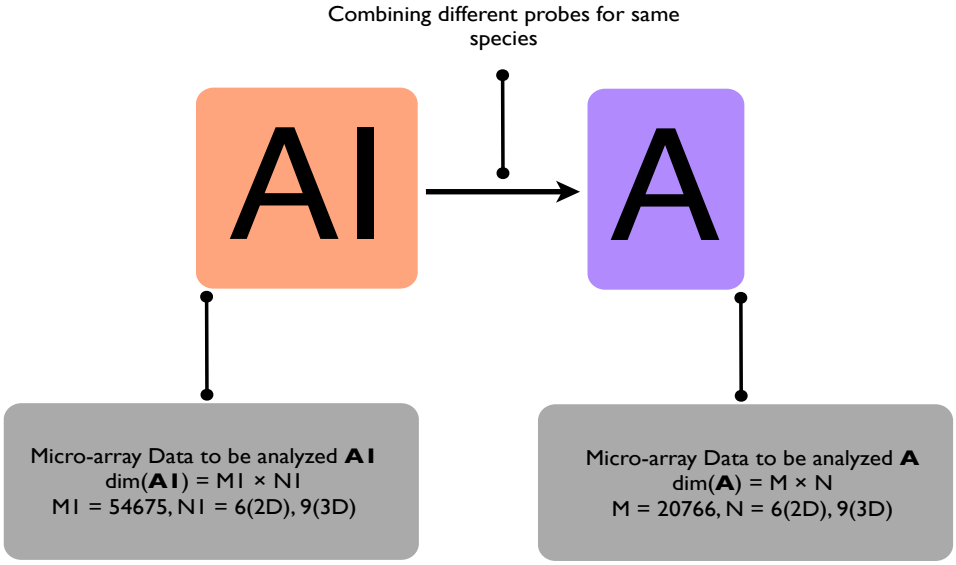


Figure 8.2: Microarray arrays have multiple probes for the same IDs. First step was to remove repetitions in order to reduce dimensions. This allowed us to consolidate multiple probes for the same species and reduce the dimensionality of the problem from 54675 to 20766.

the measurements. This allowed us to consolidate multiple probes for the same species and reduce the dimensionality of the problem from 54675 to 20766 as shown in Fig. 8.2.

Qualitative comparisons of genes in 2D and 3D experiments based on fold changes (hypoxia/normoxia) in mRNA levels showed us an interesting result (Fig. 8.3). Against our intuition that genes might be regulated in a correlated manner in 2D and 3D experiments (Quadrant 2 and 3 in Fig. 8.3), we find that the number of differentially regulated genes are more widespread in number (Quadrant 1 and 4 in Fig. 8.3).

As a matter of fact, about 1645 genes were statistically significantly differentially regulated, while only 88 genes were statistically significantly similarly

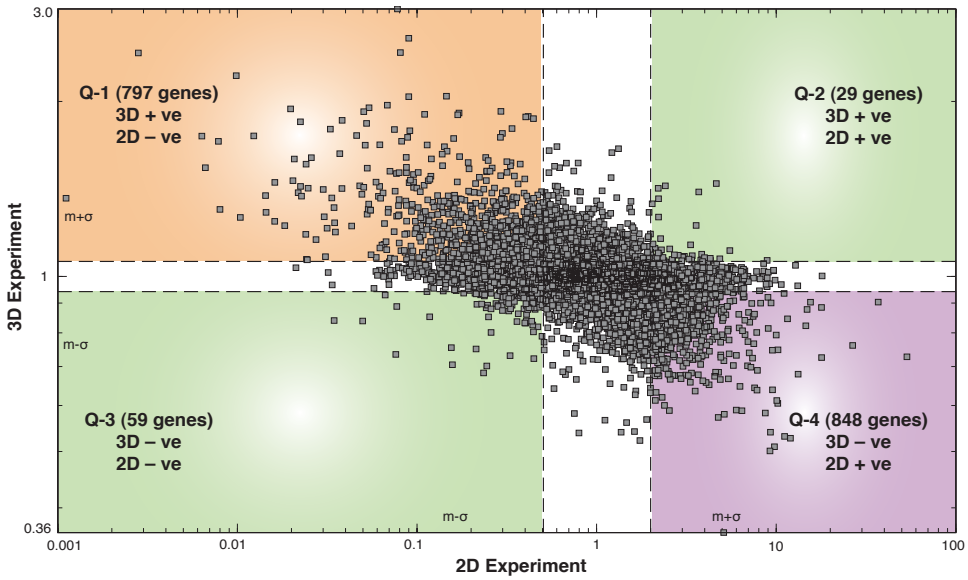


Figure 8.3: Qualitative comparisons of genes in 2D and 3D experiments based on fold changes (hypoxia/normoxia) in mRNA levels: Against our intuition that genes might be regulated in a correlated manner in 2D and 3D experiments (Quadrant 2 and 3), we find that the number of differentially regulated genes are more widespread in number (Quadrant 1 and 4). As seen in quadrant 1, MMP1/10/13, BNIP3, ANG, PDGFB are some of the key species which are differentially seen to be up-regulated in 3D as compared to 2D (down-regulated). As seen in quadrant 2, BMP5, EPHA6 are some of the key species which are consistently seen to be up-regulated in 3D as well as 2D. As seen in quadrant 3, IL11, ICAM1, IL32 are some of the key species which are consistently seen to be down-regulated in 3D as well as 2D.

regulated. In quadrant 1, MMP1/10/13, BNIP3, ANG, PDGFB are some of the key species which are differentially seen to be up-regulated in 3D as compared to 2D (down-regulated) (Fig. 8.3). In quadrant 2, BMP5, EPHA6 are some of the key species which are consistently seen to be up-regulated in 3D as well as 2D (Fig. 8.3).

In quadrant 3, IL11, ICAM1, IL32 are some of the key species which are

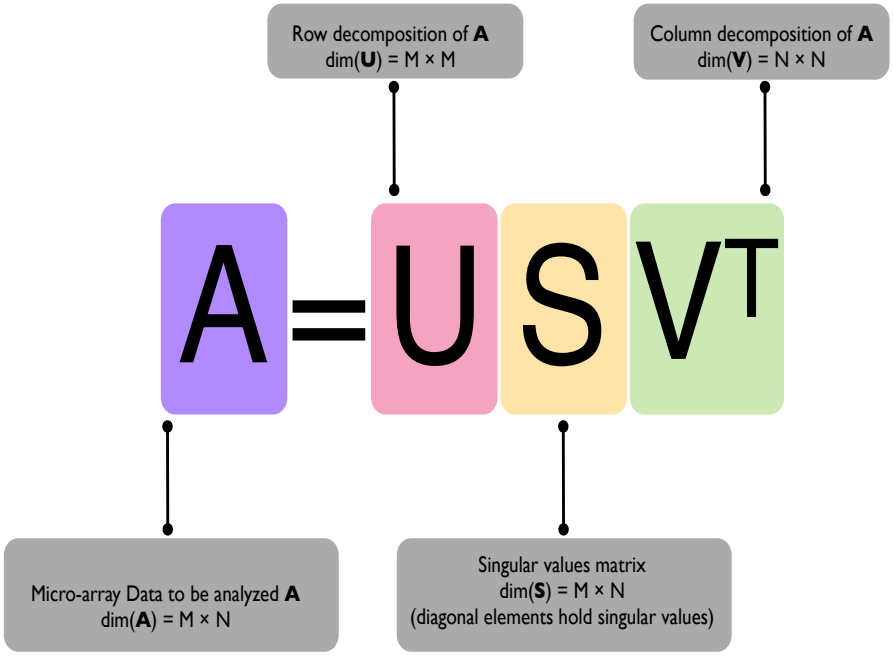


Figure 8.4: Finding natural combinations of the responses within the genes using Singular Value Decomposition (SVD) allows us to identify key regulators mediating the hypoxic response.

consistently seen to be down-regulated in 3D as well as 2D. This gives us a first insight that different regulatory components might be controlling the hypoxic conditions at different experimental protocols. For example, upregulation of MMP's in 3D differentially might suggest the importance of extracellular matrix and its interactions playing a key role in 3D culture conditions. A similar thing actually has been suggested in experiments before as well [130, 299].

For further analysis, we used Singular Value Decomposition (SVD) to find natural combinations of the responses within the genes. This allowed us to identify key regulators mediating the hypoxic response under different experimental conditions (Fig. 8.4). For this analysis we again used the average fold changes (hypoxia versus normoxia) as the matrix to decompose.

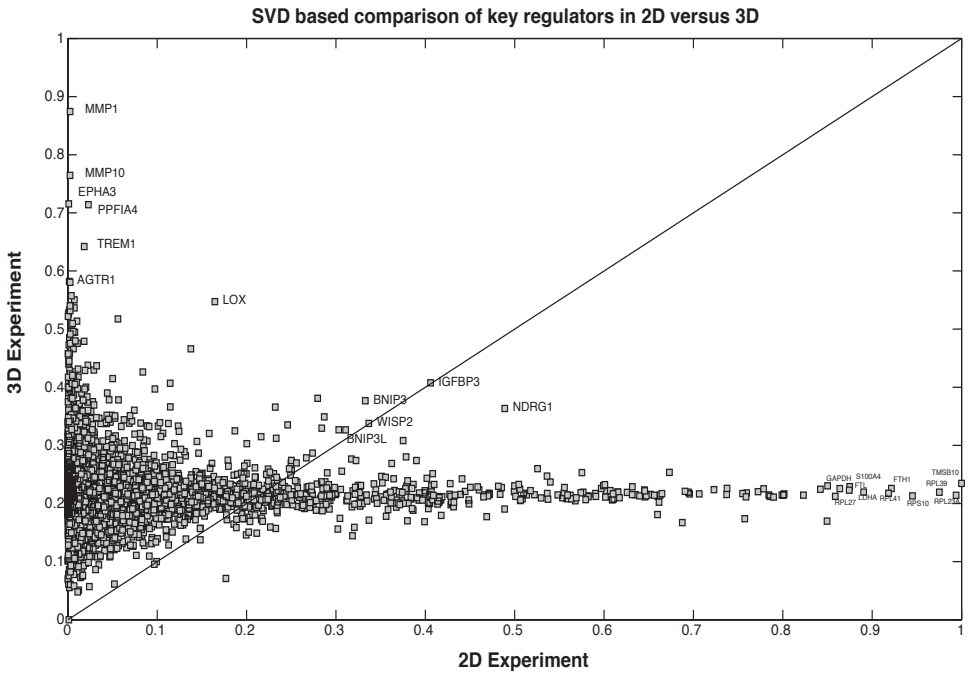


Figure 8.5: Qualitative comparisons of genes in 2D and 3D experiments based on SVD based importance: Using SVD analysis we can see the differential importance of ECM interactions in 3D as compared to 2D experiments.

The components were ranked (0-1) based on the magnitudes of the eigenvalues of the first principle component. Differentially over-expressed components like MMPs were seen to be important in case of 3D even via SVD (Fig. 8.5). For the case of 2D experiments, we saw that internal cellular metabolism components were the most important via SVD (Fig. 8.5). Interestingly, we saw that certain components were equally important (marked by on or near the diagonal line in Fig. 8.5). These included hypoxia induced death regulators like BNIP3, BNIP3L.

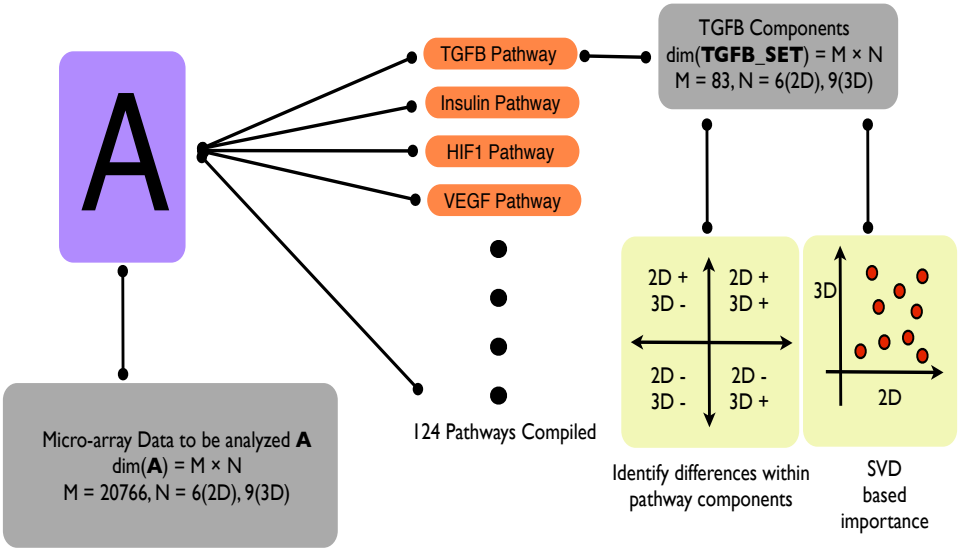
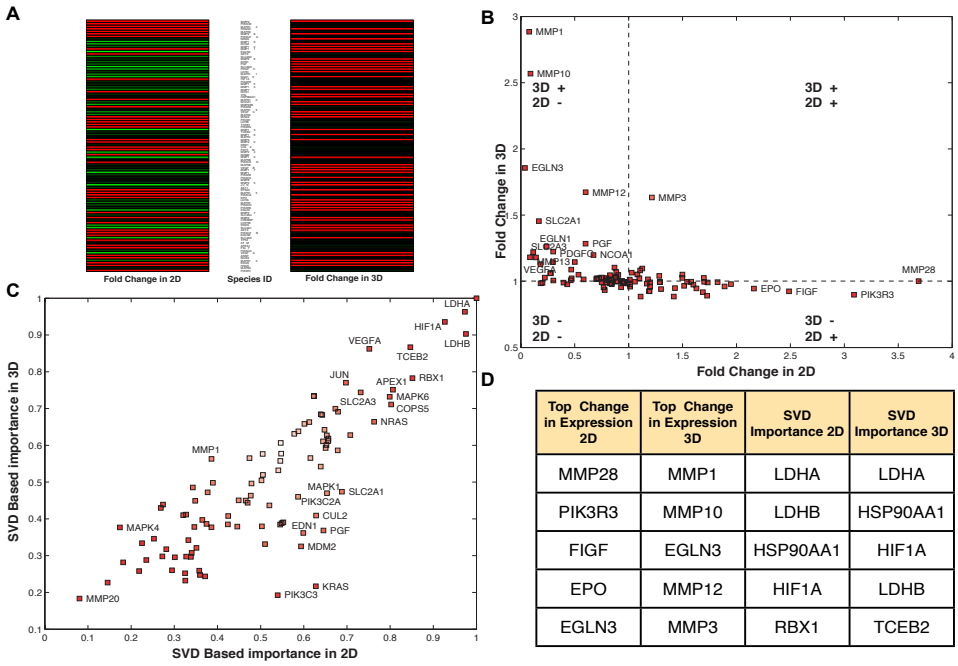


Figure 8.6: Matching the Probe IDs to the Gene names, we curated a set of 124 pathways with their corresponding participating components to investigate the differences and critical components within each pathway. For each pathway, this allowed us to identify differences within pathway components under different cases (2D and 3D) and use SVD to mine the key regulators as well. This analysis supplemented what we learned from the pathway analysis using Ingenuity Pathway Analysis.

8.3.2 Pathway Analysis of the Microarray Data

We created an in house platform to analyze microarray data and overlay the quantitative information on pre-curated pathways to identify malfunctions within the signaling pathways. For this purpose, we curated ~ 124 signaling pathways from literature and IPA (Ingenuity® Systems, www.ingenuity.com). For each pathway, this allowed us to identify differences within pathway components under different cases (2D and 3D) and use SVD to mine the key regulators as well specific to each pathway (Fig. 8.6).

For each pathway, in order to visualize the behavior we looked at the



Top Change in Expression 2D	Top Change in Expression 3D	SVD Importance 2D	SVD Importance 3D
MMP28	MMP1	LDHA	LDHA
PIK3R3	MMP10	LDHB	HSP90AA1
FIGF	EGLN3	HSP90AA1	HIF1A
EPO	MMP12	HIF1A	LDHB
EGLN3	MMP3	RBX1	TCEB2

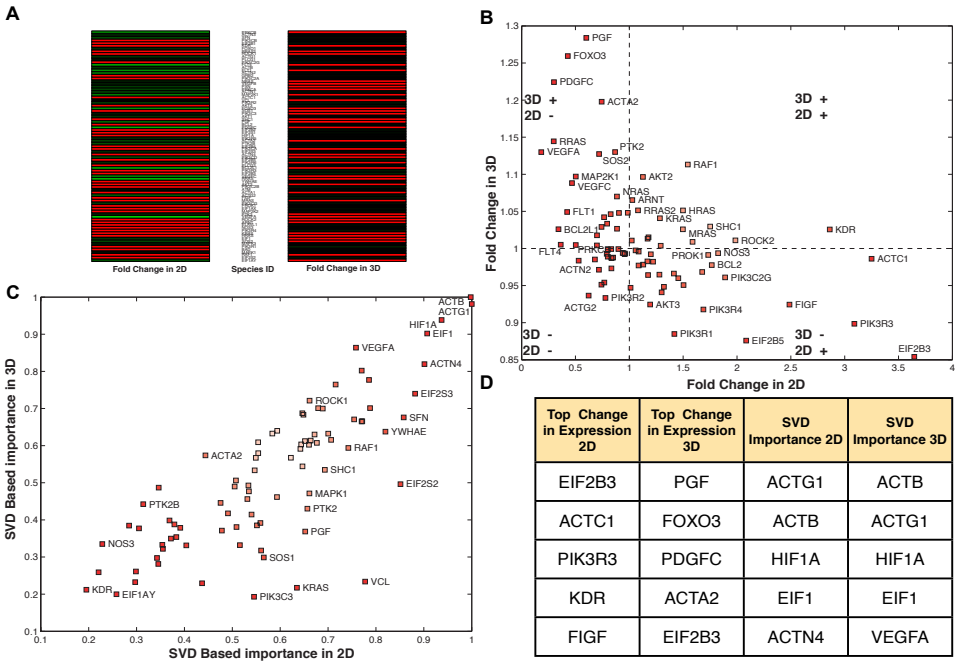
Figure 8.7: Performance of HIF1 Signaling Cascade: A) Heatmap of the components involved in the pathway. Left column - Fold change in mRNA levels in 2D experiment. Right column - Fold change in mRNA levels in 3D experiment. Color: Black - no change, Red > 1 and Green < 1. B) Cross plot of fold changes of mRNA levels in 2D (x axis) vs 3D (y axis). Each point indicates the mean of the ratio across multiple measurement sets. The quadrants of the plot indicate the similarly or differentially regulated components. C) Cross plot of importance based on the first principle direction using SVD analysis 2D (x axis) vs 3D (y axis). Values range from 0 (less important) - 1 (more important). D) Tabulation of the 5 most important components for the corresponding pathway (using absolute change in expression levels and SVD based importance). We see that MMP1/10/12 and VEGFA are some of the components differentially regulated in 2D (fold change < 1) and 3D (fold change > 1). Interestingly, HIF1α is implicated via SVD analysis to equally contribute in both cases.

heatmap of the components involved in the pathway (Color: Black - no change, Red > 1 and Green < 1). In order to quantitatively compare the cases, we looked at the cross plot of fold changes of mRNA levels in 2D (x axis) vs 3D (y axis). Each point indicated the mean of the fold change (hypoxia versus normoxia) across multiple measurement sets. The quadrants of the plot indicated the similarly or differentially regulated components. We further used SVD to investigate the key regulators at the pathway levels as well. This was seen as a cross plot of importance based on the first principle direction using SVD analysis 2D (x axis) vs 3D (y axis). Values ranged from 0 (less important) - 1 (more important). Ultimately, we tried to narrow down the analysis to identification of the top 5 components within each pathway via each analysis (as in Fig. 8.7-8.11).

In case of HIF1 signaling cascade, we saw that MMP1/10/12 and VEGFA were some of the components differentially regulated in 2D (fold change < 1) and 3D (fold change > 1). Interestingly, HIF1α was implicated via SVD analysis to equally (equal importance based on SVD ranking) contribute in both cases (Fig. 8.7).

In case of VEGF signaling cascade, we saw that VEGFA/C, PDGFC, PGF were some of the components differentially regulated in 2D (fold change < 1) and 3D (fold change > 1) (Fig. 8.8). Interestingly BCL2 was differentially upregulated in 2D (fold change > 1) suggesting that the cells were more pro-survival in 2D or struggling to stay alive in 2D as compared to 3D. In case of IL-8 signaling cascade, we saw that several components like CCND3, BAX, HBEGF were differentially regulated (Fig. 8.9). Interestingly, CXCL1 was downregulated in both cases (greater magnitude change in 3D).

With respect to regulators of angiogenesis, we saw that SDC1/2 were differ-

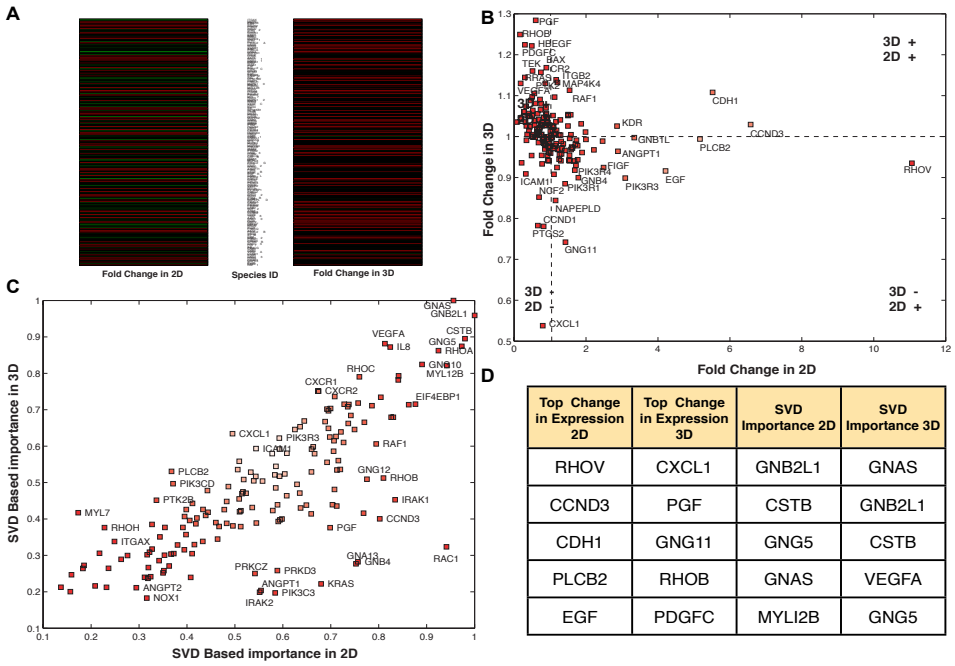


Top Change in Expression 2D	Top Change in Expression 3D	SVD Importance 2D	SVD Importance 3D
EIF2B3	PGF	ACTG1	ACTB
ACTC1	FOXO3	ACTB	ACTG1
PIK3R3	PDGFC	HIF1A	HIF1A
KDR	ACTA2	EIF1	EIF1
FIGF	EIF2B3	ACTN4	VEGFA

Figure 8.8: Performance of VEGF Signaling Cascade: We see that VEGFA/C, PDGFC, PGF are some of the components differentially regulated in 2D (fold change < 1) and 3D (fold change > 1). Interestingly BCL2 is differentially upregulated in 2D (fold change > 1) suggesting that the cells are more pro-survival in 2D or struggling to stay alive in 2D as compared to 3D.

entially up-regulated in case of 3D (Fig. 8.10). SDC1/2 is an integral membrane protein and participates in cell proliferation, cell migration and cell-matrix interactions via its receptor for extracellular matrix proteins. Recently it has been shown that SDC1 regulates αvβ3 and αvβ5 integrin activation during angiogenesis [28]. We saw an interesting behavior within components in the TGFβ signaling cascade, for example TGFβ1 was upregulated in 3D (downregulated in 2D) where as the opposite was seen in case of TGFβR2 (Fig. 8.10).

To investigate this further, we looked at the behavior of the TGFβ signaling pathway (Fig. 8.11). In 3D, we saw that BMP2, SMAD9/7, TGFβ1 were differentially upregulated while SMAD1/2/6, TGFβR2/3, BMPR1A/B were dif-



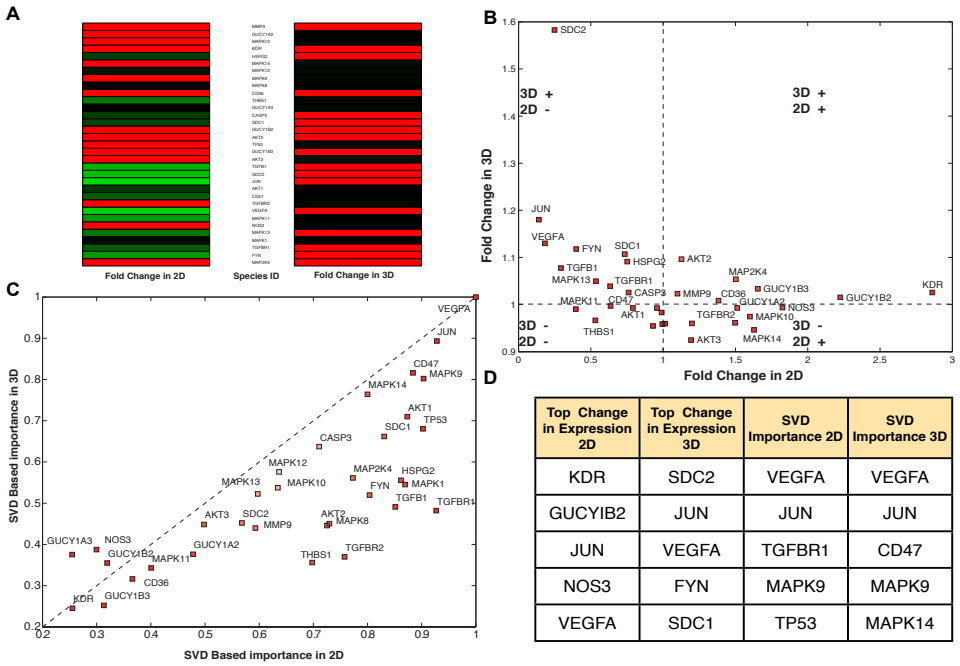
Top Change in Expression 2D	Top Change in Expression 3D	SVD Importance 2D	SVD Importance 3D
RHOV	CXCL1	GNB2L1	GNAS
CCND3	PGF	CSTB	GNB2L1
CDH1	GNG11	GNG5	CSTB
PLCB2	RHOB	GNAS	VEGFA
EGF	PDGFC	MYLI2B	GNG5

Figure 8.9: Performance of IL-8 Signaling Cascade: We see that several components like CCND3, BAX, HBEGF are differentially regulated. Interestingly, CXCL1 is downregulated in both cases (greater magnitude change in 3D).

ferentially upregulated in 2D. However, BMPR2 was upregulated in both the cases. This might suggest differential operational paradigms within the TGFβ signaling cascades operating under different experimental protocols.

8.4 Discussion

In the current study, we used a *top-down* strategy to investigate cellular behavior in case of hypoxia induced tumor growth and angiogenesis. We created an in house platform to analyze microarray data and overlay the quantitative information on pre-curated pathways to identify malfunctions within the cells. We

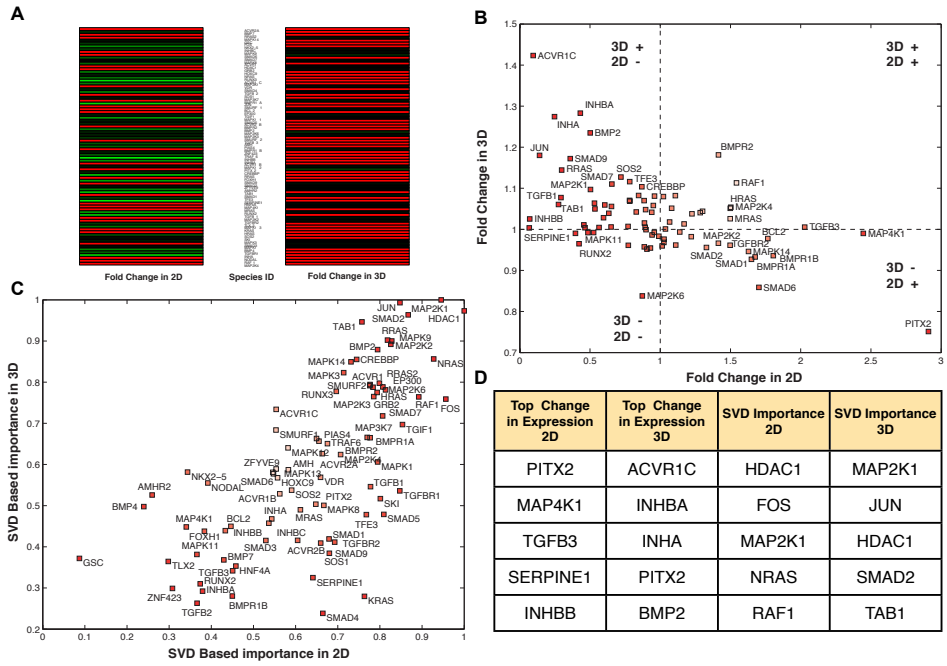


Top Change in Expression 2D	Top Change in Expression 3D	SVD Importance 2D	SVD Importance 3D
KDR	SDC2	VEGFA	VEGFA
GUCYIB2	JUN	JUN	JUN
JUN	VEGFA	TGFBR1	CD47
NOS3	FYN	MAPK9	MAPK9
VEGFA	SDC1	TP53	MAPK14

Figure 8.10: Performance of Inhibition of Angiogenesis Signaling Cascade: We see that SDC1/2 are differentially upregulated in case of 3D. We see an interesting behavior within components in the TGFβ signaling cascade, for example TGFβ1 is upregulated in 3D (downregulated in 2D) where as the opposite was seen in case of TGFβR2.

coupled this to mathematical analysis techniques like singular value decomposition (SVD) to identify key mediators of the response.

Using these strategies, we identified differential regulators of the cellular behavior in 2D versus 3D experimental protocols. We identified about 1645 genes which were statistically significantly differentially regulated, while only 88 genes were statistically significantly similarly regulated. The differentially upregulated genes in case of 3D included those interacting/influencing the cellular interaction with the extracellular matrix. This is in line with the importance of integrin and ECM interactions regulating the tumor growth and angiogenesis [153, 93]. Zooming on to the level of pathways, we identified differential behav-



Top Change in Expression 2D	Top Change in Expression 3D	SVD Importance 2D	SVD Importance 3D
PITX2	ACVR1C	HDAC1	MAP2K1
MAP4K1	INHBA	FOS	JUN
TGFB3	INHA	MAP2K1	HDAC1
SERPINE1	PITX2	NRAS	SMAD2
INHBB	BMP2	RAF1	TAB1

Figure 8.11: Performance of TGFβ Signaling Cascade: In 3D, we saw that BMP2, SMAD9/7, TGFβ1 were differentially upregulated while SMAD1/2/6, TGFβR2/3, BMPR1A/B were differentially upregulated in 2D. However, BMPR2 was upregulated in both the cases.

iors and operational paradigms of key signaling pathways within the hypoxic response. For example, in case of TGFβ signaling different modes of operation involving BMP2 was seen in 3D as compared to 2D. Similarly, VEGFa and IL-8 levels are differentially regulated in 2D and 3D, suggesting different operational modes for both as well. At the same time, both in 2D and 3D we saw that HIF1α was equally important, suggesting some basic similarities as well.

While this form of analysis provided key insights for future experimental directions (which are being pursued by our collaborators), we specifically used this insight to further improve the proof-of-concept physiochemical models of these disease states. Our first generation hypoxia model was constructed by

using data and connectivity from mostly 2D experiments (as discussed in the previous chapters). The importance of ECM interactions, Integrin signaling and differential regulation of TGFβ signaling to name a few, clearly stand out as important components to add to the next generation model. This study hopefully shall be able to bridge the gap and further our understanding of differences and similarities between 2D and 3D experimental protocols. Besides, that this overall strategy in itself has immense scope of development, ranging from identifying key regulators to forming proof-of-concept physiochemical models.

8.5 Materials and Methods

8.5.1 2-D Studies and 3-D Alginate Tumor Models

Experiments were carried out by our collaborators in Biomedical Engineering and Weill Cornell Medical School. For 2D, OSCC3's were cultured at 37°C, 5% CO_2 and either 1 or 17% O_2 in controlled atmosphere incubators (Thermo Fisher Scientific, Inc., Waltham, MA) for 6 days. Alginate hydrogels (3% [w/v]) were used as 3-D cell scaffolds. Alginate (Protanal LF; FMC Biopolymer, Philadelphia, PA) was dissolved in serum-free DMEM, and OSCC3 cells were suspended in this dissolved alginate at a density of 20×10^6 cell/mL. Cell-seeded alginate gels were cast into 4 mm diameter, 200 μm thick cylinders using a plexiglass mold, and cross-linked with 100 mM $CaCl_2$ for 15 min. Discs were then cultured in 24-well plates, with a single disc and 500 μL culture media (DMEM + 10% FBS + 1% penicillin/streptomycin) per well. Cells were cultured at 37°C, 5% CO_2 and either 1 or 17% O_2 in controlled atmosphere incubators (Thermo

Fisher Scientific, Inc., Waltham, MA). Media was changed daily.

8.5.2 RNA Target Preparation/Affymetrix Microarray Analysis

Experiments were carried out by our collaborators in Weill Cornell Medical School. Total RNA was extracted from the samples after exposure to 6 days of normoxia/hypoxia. All relevant protocols involving refining and quality assessment was done as well. Subsequently, total RNA was processed using a Affymetrix GeneChip.

8.5.3 Evaluation of Affymetrix GeneChip Data

The hybridizations were normalized by using the RMA (robust multichip averaging) method to obtain summary expression values for each probe set [226]. Gene expression levels were analyzed on a logarithmic scale.

CHAPTER 9

CONCLUSIONS AND FUTURE DIRECTIONS

Eukaryotes have developed evolutionarily conserved mechanisms to respond to diverse ranges of internal and external perturbations e.g., changes in oxygen/nutrient levels, temperature oscillations, protein folding load, viral/bacterial attacks and graft implants. Alterations and malfunctions within these core regulatory architectures provide a peek into shifts towards cancerous and disease states. In my research, I have used a combination of *bottom-up* analysis (signaling networks based analysis) and a *top-down* analysis (using microarray/high throughput experimental data) to investigate stress responses in eukaryotes. Even with the constraints in the form of biological know-how and materials and methods, we were able to generate experimentally verifiable hypothesis in several aspects of the eukaryotic stress responses. However, we need to address some open areas as we move further in this field e.g., addressing structural completeness of models, structural analysis of models, model reduction strategies, alternative model solution techniques. A better understanding of cellular behavior and the underlying regulatory architecture will be obtained from an integrated approach that simultaneously incorporates the individual and contextual properties of all constituents in complex cellular networks [234].

9.1 Structural completeness for next generation models

In the *bottom-up* analysis scheme, structural uncertainty in the form of missing species/connections (either or both) are a huge issue. For example, in UPR while we did an extensive search of the literature to formulate the model, we

were missing certain key structural aspects of UPR which could provide a more comprehensive analysis for further generations of the model. One key missing aspect is the negative regulation of the three ER-stress transducers. Given PERK's central role in translation attenuation, cells have evolved multiple axes to regulate PERK activity. First, the cytosolic kinase domain of PERK can be inhibited by the action of the DNAJ family member P58^{IPK}. P58^{IPK} was initially discovered as an inhibitor of the eIF2α protein kinase PKR [292]. P58^{IPK}, whose expression is induced following ATF6 activation, binds to the cytosolic kinase domain of PERK, inhibiting its activity [599, 558]. Inhibition of PERK kinase activity relieves eIF2α phosphorylation, thereby removing the translational block. Interestingly, P58^{IPK} expression occurs several hours after PERK activation and eIF2α phosphorylation. Thus, P58^{IPK} induction may mark the end of UPR adaptation, and the beginning of the alarm/apoptosis phase of the response [515]. Second, PERK induces a negative feedback loop, through its downstream effector CHOP, involving the direct de-phosphorylation of eIF2α. CHOP induces the expression of GADD34 which, in conjunction with protein phosphatase 1 (PP1), assembles into a phosphatase which dephosphorylates the S51 residue of eIF2α [384]. GADD34 is a member of the GADD family of genes which are induced by DNA damage and a variety of other cellular stresses [621]. The GADD34 binding partner in this complex appears to be responsible for PP1α recognition and targeting of the phosphatase complex to the ER. Association between GADD34 and PP1 is encoded by a C-terminal canonical PP1 binding motif, KVRF, while approximately 180 residues, near the N-terminus of GADD34, appear to be responsible for ER localization [56]. Currently, little is known about deactivation of ATF6. Recently, XBP1u, the unspliced form of XBP1, has been implicated as a negative regulator for ATF6 [615]. Following, the induction of ER stress, two

versions of XBP1 exist: XBP1u and sXBP1 [615]. In the recovery phase following ER stress, high levels of XBP1u may play a dual role. First, XBP1u binds sXBP1, promoting complex degradation [614, 541]. Second, XBP1u can bind ATF6α rendering it more prone to proteasomal degradation [615]. Taken together, these two steps may slow the transcription of ER chaperones and ERAD components during the recovery phase following ER stress. IRE1α activity is regulated by several proteins, including tyrosine phosphatase 1B (PTP-1B), ASK1-interactive protein 1 (AIP1) and members of the Bcl2 protein family. PTP-1B has been implicated in a number of IRE1α signaling events. The absence of PTP-1B reduced IRE1α dependent JNK activation, XBP1 splicing and EDEM transcription in immortalized and primary mouse embryonic fibroblasts [171]. However, no physical interaction between IRE1α and PTP-1B was established. On the other hand, AIP1 physically interacts with both TRAF2 and IRE1α, suggesting a model in which AIP1 facilitates IRE1α dimerization and activation [321]. The C-terminal period-like domain (PER) of AIF1 binds the N-terminal RING finger domain of TRAF2, followed by ASK1-JNK signaling [623]. Thus, based on these findings, Luo *et al.* postulated that AIF1 may be directly involved in the IRE1α-TRAF2 complex and its activation of the ASK1-JNK signaling axis [321]. This hypothesis was validated in AIP1-KO mouse studies; AIP1-knockout mouse embryonic fibroblasts and vascular endothelial cells showed significant reductions in ER-stress induced ASK1-JNK activation that was rescued in AIP1 knock-in cells [321]. IRE1α has also been shown to directly interact with Bcl-2 family members Bax and Bak. Hetz *et al.* showed that Bax and Bak complex with the cytosolic domain of IRE1α and modulate IRE1α signaling [207]. Bax and Bak double knockout mice failed to signal through the IRE1α UPR branch following tunicamycin-induced ER stress; however, PERK signaling markers, e.g., eIF2α

phosphorylation, responded normally [207]. This pro-activation role of Bak and Bax may be modulated by one of the few negative regulators of IRE1α activity, Bax inhibitor 1 (BI-1). BI-1 is an anti-apoptotic protein that enhances cell survival following several intrinsic death stimuli [593]. Bailly-Maitre *et al.* were the first to suggest that BI-1 may downregulate IRE1α and possibly ATF6 activity [21]. BI-1 deficient mice displayed increased XBP1s and enhanced JNK activity in the liver and kidney, while eIF2α phosphorylation remained normal under ER-stress conditions [21]. Lisbona *et al.* later showed that BI-1 directly interacts with the cytosolic domain of IRE1α, inhibiting its endoribonuclease activity [307]. Interestingly, BI-1 interacts with several members of the Bcl2 protein family e.g., Bcl2 and Bcl-X_L, even though it has no homology [593]. Members of the HSP family of proteins have also been shown to regulate IRE1α. For example, HSP90 interacts with the cytosolic domain of IRE1α, potentially protecting it from degradation by the proteasome [332]. HSP72 interaction with the cytosolic IRE1α domain has also recently been shown to enhance IRE1α endoribonuclease activity [175]. Taken together, these modes of IRE1α regulation with the exception of B1-1, largely promote or enhance IRE1α signaling. Given the importance of CHOP in regulation of Bcl2, it is vital to establish the exact connectivity. However, while CHOP expression is negatively correlated with Bcl2 levels, there is no CHOP binding site in the *bcl2* promoter [345]. McCullough *et al.* have suggested that the bZIP domain of CHOP could act with other bZIP transcription factors to regulate *bcl2* expression [345]. Thus, it's likely that the connection between CHOP expression and apoptosis is more complex than simple down-regulation of Bcl2 expression. These missing structural connections shall allow us to establish a detailed model and extract more relevant insights into manipulating UPR.

Similarly in case of translation initiation, model connectivity was assembled from an extensive literature review. However, several potentially important signaling mechanisms were not included. First, we should revisit the role of PRAS40. Currently, PRAS40 acts as a cofactor that aids in pAkt(Ser473)-mediated activation of mTORC1. Sancak *et al* suggested that PRAS40 sequesters mTORC1, and only after phosphorylation by Akt does it releases from mTORC1 [469]. Other groups have also shown that mTORC1 can phosphorylate and inhibit PRAS40, thus providing a positive feedback mechanism for Akt-mediated mTORC1 activation [132, 572]. A more complete description of PRAS40 will enhance our ability to interrogate Akt dependent mTORC1 activation. Second, we need to refine the description of IRS-1 feedback. Currently, we assume a single deactivating phosphorylation event at Ser308. However, several studies have shown that IRS-1 can be phosphorylated at multiple serine sites, which are both activating and deactivating [414, 173]. Additionally, PTEN is known to dephosphorylate activated PDGF receptors and attenuate their activity, a feature not included currently [519]. A more complete description of IRS-1 phosphorylation could help define how, and under what conditions, IRS-1 regulation attenuates PI3K activation. Third, we modeled the regulation of 4E-BPx as a single phosphorylation event where phosphorylated 4E-BPx was unable to bind to eIF4E. In reality, 4E-BPx family members, such as 4E-BP1, have several phosphorylation sites [442] and the release of eIF4E is driven only after multiple conserved phosphorylation events [162]. Additionally, eIF4E can itself be phosphorylated at Ser209; while there is agreement that the phosphorylation of eIF4E does have a regulatory significance, the data is contradictory as to whether it is positive or negative [474]. Fourth, signaling downstream of mTORC1 has also been shown to mediate translation modes beyond those included in our model. eIF3 has

been identified as a scaffolding protein that recruits mTORC1 to untranslated mRNA and facilitates S6K1 and 4E-BP1 phosphorylation [214]. S6K1 can also activate eIF4B, a protein that helps eIF4A to unwind the secondary structure of untranslated mRNA [443]. Further, a recently discovered scaffold protein, SKAR, has been shown to assist S6K1 recruitment to mRNA [323]. Lastly, because of mTORC1's unique cellular role, it would be interesting to explore how other aspects of metabolism interact with insulin signaling to mediate decisions between translation, lipid synthesis or proliferation. In these studies, one could imagine constructing *in-vivo* mouse models to explore the physiological role of mTORC1 signaling in important diseases such as diabetes or cancer.

In case of EMT model, we lacked certain modules of signaling within EMT. In particular, the regulation of β-catenin through the GSK3β complex is a widely studied and intricate component of the EMT process [244]. TGFβ 3 signals through PI3K to activate AKT [329, 618]. AKT can then phosphorylate and inactivate the GSK3β complex [618]. GSK3β is an important target of both Snail and β-catenin for degradation through the ubiquitin-proteasome pathway [244]. Thus, it provides a critical negative feedback mechanism which would should be incorporated in further models. Specifically to our model, expression of Snail increases through 72 hours. In contrast, experimental data has shown that activity of Snail peaks at 24 hours. This is possibly through the regulation of GSK3β and should be included in future models. Addressing missing structural components like that of GSK3β could allow us to get a more comprehensive insight into TGFβ signaling. Given the scope of the problem and limitations imposed by structural/parametric uncertainties, we were able to investigate into the mechanistic workings of EMT and present falsifiable hypothesis relevant to manipulating EMT. While this study was limited to EMT, a similar framework could be

used to understand and investigate other similar operational modes relevant to biology.

In case of EHR, we are lacking some details in regulation of HIF-1α, which taken into account might reveal a clearer picture [25]. In the current EHR model, HIF-1α was regulated only at the protein level. We only have the constitutive expression of HIF-1α and HIF-1β. In order to get a complete picture of regulation of HIF-1α, we should include the negative regulation of HIF-1α via the regulation of PHD2 protein. PHD2 mRNA and protein have been reported to be up-regulated in case of hypoxia, leading to negative feedback [25]. Transcriptional role of HIF-1 is mediated via binding to p300 protein. CITED2/4 proteins (increased expression in hypoxia) have been identified which competitively binds to p300 thereby operating in a negative feedback loop to down-regulate HIF-1 mediated transactivation [25]. p53 has been seen to promote Mdm2-mediated ubiquitination and proteasomal degradation of HIF-1α [445]. Furthermore there exists a positive feedback to HIF-1α signaling mediated by growth factor and cytokines via Ras/Raf/MAPK and PI3-K/Akt/PKB cascades [25]. These pathways also play a role in regulating the HIF-1 signaling cascades in Normoxia. Given the role of VEGF and IL-8 signaling pathways in angiogenesis, it has rightly been targeted in several efforts in regulating tumors [591, 271, 126]. Regulation of VEGF by the humanized monoclonal antibody bevacizumab, is currently used successfully for the treatment of cancer [127]. In the current model, VEGF mediated signal initiation has been modeled via VEGFA, VEGFR1/2/3, heparin sulfate proteoglycans (hspg) and neuropilins. Given the importance of VEGF signaling, we might get a clearer and more specific picture upon inclusion of specific details regarding variants of VEGFA (VEGF-206, VEGF-165 and VEGF-183), other ligands like VEGFB/C/D/E and the specific role of

Neuropilin-1 [458]. Neuropilin-1 mRNA and protein have been seen to be over-expressed in case of human colon growth and progression [409]. Downstream effects of VEGF signaling cascade mediated by PI3K-AKT module have also not been taken into account in the current EHR model [562, 69]. Both structurally and functionally relevant, role of AP-1 and NF-$\kappa\beta$ has been identified by the model analysis to be of high importance. Rightfully so, regulation of NF-$\kappa\beta$ has been seen to be critically involved in cancer development and progression, as well as in resistance to chemotherapy and radiotherapy [27]. Thereby, the therapeutic potential and benefit of targeting NF-$\kappa\beta$ and ROS in cancer holds great promise [27, 555]. Currently, a narrow account of NF-$\kappa\beta$ regulation has been taken into account in the EHR model. It shall be interesting to incorporate the detailed complexity associated with regulation of NF-$\kappa\beta$ and then analyze the critical aspects of the EHR model [86]. In terms of NF-kB, we have not taken into account the regulation of IKK via PKC. IKK traditionally leads to Ser32 and Ser36 phosphorylation of IKBα, thereby targeting it for proteosomal degradation [86]. However in case of hypoxia another alternate route of IKBα phosphorylation not including proteosomal degradation also plays a key part [86]. In the current EHR model, we have AP-1 regulation via p38MAPK and ROS. We have not taken into account the individual dimer components, namely cFOS and cJUN. To get a comprehensive insight into the role of AP-1, we should include specific details like ERK2 mediated phosphorylation of cFOS at T232 to establish the link between ERK and AP-1 [23]. MAP kinases ERK, JNK and p38 phosphorylate and activate Elk-1 resulting in enhanced SRE-dependent cFOS expression [586]. JNK and p38 MAP kinases lead to activation and enhanced expression of cJUN [586].

One problem we face currently is the reliability of network data and how

we join them up in networks as recently talked about by Nathan Blow [45]. While POETs allows us to obtain trade-offs between competing data sets, there is a need to deal with data reliability in such complex systems. In that regard, Guimera *et. al.* recently presented a mathematical and computational framework to deal with the problem of data reliability in complex networks [174]. They were able to reliably identify both missing and spurious interactions in noisy network observations. We could use this to guide experiments, to better characterize network data sets, and to drive new discoveries.

9.2 Modeling based experimental design

Model analysis has provided valuable insights and several experimentally verifiable hypothesis to manipulate the cellular stress responses. For example, in case of EUPR, the key role of ATF4 was brought to light. We can test this experimentally by knocking out (KO) ATF4 and studying the effect on the cellular death/survival phenotype. Furthermore ATF4 KO shall also effect the temporal response of the EUPR. This can be studied using western blots for key markers such as BiP, CHOP and some apoptosis markers. In case of EHR, AP1 was identified as the key in regulating the delayed autocrine mediated amplification of the hypoxic response. Experimentally we can check this by either studying an AP1 KO system or by using MAPK inhibitors to prevent AP1 activation. We can measure the response of the system by looking at key markers like secreted VEGF, IL8 levels. In case of EMT, the balance between positive and negative regulation of SMADs and activation of AP1/SP1 was seen as a key regulatory aspect. We can test this experimentally by using AP1/SP1inhibitors or AP1/SP1 KO cells to study the effect on EMT. Measurement of key markers

like E-Cadherin, Vimentin, LEF1 shall elucidate the effect of these changes on the response.

9.3 Structural analysis of intracellular signaling networks

Signal transduction architectures frequently contain redundancy, feedback and crosstalk. These topological features ensure signal propagation is adaptable, efficient and robust. However, they also make reprogramming signal flow challenging. As we progressively understand the importance of a protein as an element in a network of protein-protein interactions having a contextual or cellular function within functional modules, there is a growing need of development of tools for structural analysis of complex networks [234].

Study of larger networks could be facilitated by deconstructing such complex networks into conceptually simple entities. In this regard, identification of network modules using structural and deterministic tools could be a step forward [408]. Girvan *et. al.* proposed a method for detecting communities, built around the idea of using centrality indices to find community boundaries [163]. Newman *et. al.* recently presented a set of algorithms for discovering community structure in networks [373]. These techniques would allow natural divisions of network nodes into densely connected subgroups which could be used to shed light on the sometimes dauntingly complex structure of networked systems. Addressing the similar problem, Yu *et. al.* recently presented a novel algorithm combining Molecular Complex Detection (MCODE) with Girvan-Newman (GN) to identify modules in Protein-Protein Interaction (PPI) networks [617]. These and others could be used as an effective, reliable

and scalable method of identifying modules in PPI networks [41].

While the above techniques can provide modular insights, to understand the structural basis of crosstalk and redundancy of a network, we need to look at paths through the network. One approach is to use extreme pathways (EPs). Extreme pathways (EPs) are systemic pathway vectors which describe steady-state flux distributions through molecular networks. While EPs tell us nothing about network dynamics, the convex combination of these vectors provides a means of understanding possible steady-state phenotypes. Thus, extreme pathways and elementary modes [482] have been used previously to understand the qualitative features of metabolic and signaling networks [511, 406, 405]. Singular value decomposition (SVD) of the stoichiometric matrices [121] and EPs could be a useful tool to characterize capabilities of such large scale networks. Papin, Price and Palsson *et. al.* have presented frameworks for this approach [407, 428, 408]. Dasika *et. al.* recently presented an alternate optimization-based framework for elucidating the input-output structure of signaling networks [90]. This approach could be used to exhaustively identify all input-output connections implied by the signaling network structure. Identification of paths through the network could attest for structural completeness, identify missing links and also support basis for crosstalk and redundancy within the network.

9.4 Intracellular signaling model reduction strategies

As we make leaps and bounds of progress in unraveling the underlying biological mechanisms, modeling such signal transduction networks are affected by the problem of combinatorial complexity [211]. For example, receptors and scaf-

fold proteins possess a large number of distinct binding domains leading to the formation of large multiprotein signaling complexes. Owing to this combinatorial reason the number of distinguishable species grows exponentially. Even by including only a limited number of components and binding domains the resulting models are very large and hardly manageable [80, 266, 211, 81]. Most of the current modeling strategies evade this problem by assuming reduced structures which focus on restricted number of species and reactions e.g., my current work. However, we need to incorporate more mechanistic and component details to get a clearer picture.

Faeder *et. al.* showed that only a relatively small part of the network is actually active in any state and eliminating the other low concentration species, a fairly reduced model could be built [120]. However, perturbations in parameters over such reduced networks led to large approximation errors. Conzelmann *et. al.* have presented several alternative methods to reduce models of combinatorial reaction networks [80, 81]. Kremling *et. al.* recently reviewed several other strategies for model reduction [273]. As we move forward with increased model size and complexity, we could explore some of these techniques.

9.5 Alternate modeling strategies for biological systems

As more experimental information is available thanks to improved technology, it is becoming imperative to further modeling techniques to make full use of the information. Gillespie *et. al.* have presented quite a few stochastic methodologies for systems biology [159]. Some of these methodologies currently in use are explicit tau-leaping and the slow-scale stochastic simulation algorithm [62]. Re-

cently *petri-nets* based modeling of biological networks is gaining importance as an alternative to model complex systems [71, 569]. Heiner *et. al.* have used petri nets extensively to study metabolic systems and also to infer structural information about networks [569, 196]. Ruths *et. al.* used a petri net execution strategy for characterizing the dynamics of signal flow through a signaling network using token distribution and sampling [460]. The hallmarks of this strategy is improved speed and ease to obtain insights into the trends of species activity based solely on network's connectivity. Sandman *et. al.* recently presented a stochastic discrete-time Markov chain method of modeling complex biological systems [471]. This method was claimed to be more efficient computationally to the Gillespie algorithm. Talcott *et. al.* have explored another approach called *Pathway Logic* to model causal networks of biomolecular interactions in a logical framework at multiple scales [197, 522]. Another aspect of modeling such biological systems is the possibility to parallelize the simulations. In this regard, Zhou *et. al.* recently presented a graphics processing units (GPU) based strategies of modeling biological systems [634]. Exploring these strategies might be useful as we move further ahead in terms of model size and complexity.

9.6 Multiscale modeling techniques

Multiscale modeling strategies involving extrapolation of the molecular signaling events to tissue-level/organ-level phenotypes is an emerging area of research. We presented a signaling assisted multiscale modeling technique (SAMM) to study systems across multiple length scales. Key highlights of SAMM included use of detailed cellular models to regulate cellular decisions, use of an ensemble of models to incorporate population based behavior of cells,

coupling intracellular markers and probabilistic rules to monitor cellular fate and cellular motility and a hybrid approach to improve speed of solving of external and internal mass-balances. SAMM allowed us the flexibility to include multiple cell types, freedom of seeding cells with specific internal cellular machinery within complex geometries, regulate multiple cellular decisions using internal and secreted cellular markers coupled with probabilistic functions and improved solving speed. This platform allowed us to investigate complex multiscale phenomenon e.g., tumor growth, angiogenesis.

However there are several avenues of improvement in the current strategy. For accurate predictions and analysis, we need more detailed/comprehensive intracellular cellular signaling models and additional cell type models e.g., tip cell, stalk cell models. Using different cell type models with inherent differences in machinery might provide a clearer picture regarding heterogeneity and elucidate spatio-temporal variations along different segments of growth. Having detailed models will further create an avenue to improve computational capabilities for the same.

There could be a diverse range of factors influencing the morphology and dynamics of growth e.g., concentrations of ECM proteins, gradients of growth factors, neighboring cellular densities, nutrient levels, physical stresses. SAMM allows us the flexibility to study the influence of these external factors on growth morphologies. Just by incorporating neighboring cellular density information we could transition from a bulk growth (as seen in tumors [55]) to well formed dendritic growth with/without a bias towards corners and edges (as seen in 3D cultures [371]). In reality there must be more at play. Looking forward, using SAMM we can reverse engineer a particular growth morphology by sam-

pling the probability function landscape. This shall allow us to test the contribution and implications of different extracellular factors in determining the nature of growth. Another area which we barely touched upon in the current study was use of intracellular markers to determine cellular phenotype. For example, we saw differences in growth when we based the proliferative moves based on intracellular levels of CYCLIND and ppERK. This begs further investigation with regards to choice of individual or combinatorial use of intracellular marker levels to model phenotypic decisions. Using SAMM, we can identify the relevance/validity of these choices.

APPENDIX A

HYPOXIA - SUPPLEMENTARY MATERIALS

A.1 Molecular mechanisms involved in EHR

A.1.1 Oxygen sensing.

Cells use several mechanisms to sense and respond to changes in external oxygen levels. Oxygen in the microenvironment is sensed by hypoxia inducible factor 1 α (HIF1α) and the generation of reactive oxygen species (ROS) [300, 355]. HIF1α mediates the initial phase of the angiogenic program by forming a transcriptionally active complex with HIF1β and co-activators such as p300. The stability of the HIF1α subunit is oxygen dependent [429]. In normoxic conditions, hydroxylation at two prolyl residues (P402 and P564) by PHD proteins promotes the association of HIF1α with the Von Hippel-Lindau (VHL) E3 ubiquitin ligase and subsequently leads to degradation. An additional hydroxylation site at N803 near the C-terminus of HIF1α is regulated by the asparaginyl hydroxylase FIH. Hydroxylation at N803 does not influence stability; rather, it blocks the interaction of the HIF1α C-terminal domain with transcriptional co-activators such as p300. Activated HIF1 up-regulates the expression of many factors including VEGF and Interleukin-8 (IL-8) [580]. On the other hand, the second oxygen dependent signaling axis, ROS, promotes nuclear factor κB (NF-$\kappa\beta$) activation [355]. NF-$\kappa\beta$ also regulates both VEGF and IL-8 expression [355, 580]. The exact relationship between ROS and NF-$\kappa\beta$ activation is unclear. However, ROS has been hypothesized to activate a serine kinase which in-turn phopshorylates the N-terminal serine residues (S32/S36) on IKK; potential mechanisms

reviewed in [131]. Secreted VEGF and IL-8 signals can then amplify the hypoxic response and induce proliferative signals via both autocrine or paracrine signaling mechanisms through the VEGFR and CXCR receptor families. HIF1 and NF-κB also play a key role in cellular fate by regulating the pro-death and pro-survival factors BNIP3 and BCL2.

A.1.2 VEGF and IL-8 mediated growth factor signaling.

HIF1 transcriptional activity results in the production of growth factors PLGF and VEGFA. PLGF and VEGFA are transported outside of the cell and associate with cell surface receptor VEGFR1, a decoy receptor that does not result in any cytoplasmic signaling in our model. However in recent experiments it was shown that VEGFR1 signals through the eNOS pathway leading to implications in the VEGF signaling pathway [535]. VEGFA can also associate with VEGFR2 and VEGFR3. Binding of VEGFA to VEGFR2 results in receptor phosphorylation and the activation of numerous signaling pathways. Phosphorylated VEGFR2 binds with phospholipase C γ (PLCγ) which in turn binds and activates protein kinase C (PKC). Activated PKC phosphorylates IkB$\alpha$and allows NF-κB (p50-p65) mediated transcriptional regulation. Activated PKC also leads to Ras-GDP to Ras-GTP conversion, which leads to the activation of the traditional RAF-MEK-ERK MAPK cascade via RAF phosphorylation. Double-phosphorylated ERK leads to cPLA2 phosphorylation which results in in further PKC activation. Phosphorylated VEGFR2 can also associate with focal adhesion kinase(FAK), Shb and p38MAPK to activate Paxillin, PI3K and HSP27, respectively. Hypoxia induced transcriptional activity results in increased levels of CXCR1 expression and IL8 secretion. Extracellular IL8 binds to CXCR1

resulting in association of cytoplasmic G-coupled proteins and downstream signaling. One route of Activated CXCR1 signaling occurs through PI3K. PI3K subsequently activates src and FAK. CXCR1 can also signal through Janus Kinase(JAK) or p38MAPK.

A.1.3 Transcription factor regulation.

In addition to HIF-1α, we modeled the expression and regulation of four other transcription factors: AP-1, ATF2, MEF2 and NF-$\kappa\beta$. AP-1 expression was dependent upon both ATF2 and MEF2. Both ATF2 and MEF2 were constitutively expressed and regulated by phosphorylation at T69/T71/T73 for ATF2 and T312/T319 for MEF2 by p38MAPK family members. We modeled a single activating phosphorylation site on AP-1 (S63/S73), which was regulated by ppERK (phosphorylated ERK) which is formed downstream of both IL-8 and VEGFA signaling cascades. NF-$\kappa\beta$ regulation was more complex; NF-$\kappa\beta$ was assumed to be sequestered by a lumped pool of IκB family members. The regulation of this pool occurred by phosphorylation at S32/S36 and Y42 on IκB; S32/S36 was phosphorylated by PKC while Y42 was regulated by an unknown ROS-dependent kinase. Lastly, protein markers downstream of these transcription factors were used to characterize proliferation, survival and death phenotypes. AP-1 dependent Cyclin D expression was used as a marker of proliferation. NF-$\kappa\beta$ dependent Bcl2 expression was used to characterize cell survival while HIF-1α dependent BNIP3 expression was used as a cell death marker. In addition to these markers, the expression of VEGFA, IL-8 and their respective surface receptors VEGFR2 and CXCR1/2 was also modeled.

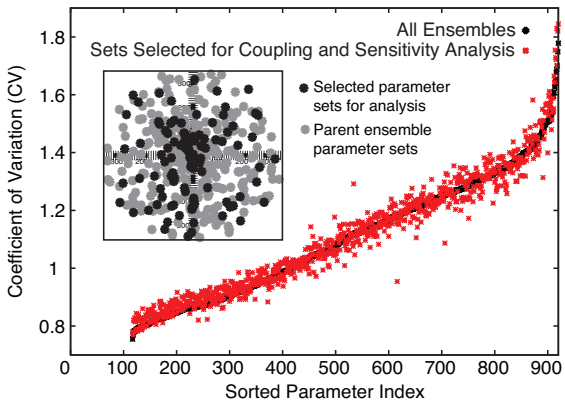


Figure A.1: Average coefficient of variation (CV) for the parent ensembles lie in the range of 0.8-1.8 (marked by the black line). 100 parameter sets were selected based on CV (CV plot in red), to be used in further analysis. Radial plot of the second norm of the parameter sets selected show diversity in generation and selection of parameter sets in the ensemble (Inset).

A.1.4 Transcription, translation and transport.

Genes for all species were constitutively expressed in the nucleus and were transcribed at basal levels by RNA Polymerase(RNAP). The binding of transcription factors to the promotor regions of genes resulted in increased transcriptional activity by stabilizing RNAP-gene binding. Transcribed mRNA was transported into the cytoplasm where it was bound and translated by assembled ribosomes. Several cytosolic components (growth factors and cytokines) can be transported out of the cytoplasm and into the extracellular matrix to perform autocrine signaling.

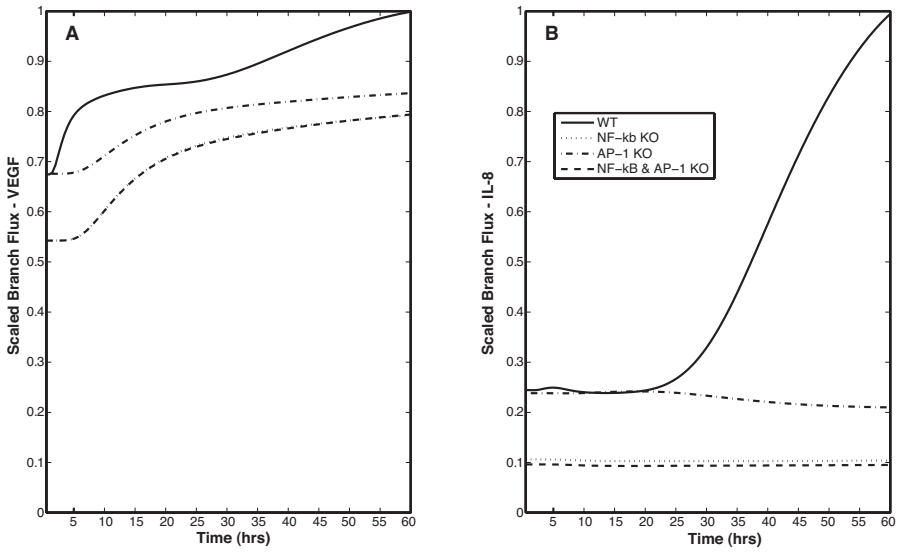


Figure A.2: KO effect on VEGF and IL-8: AP-1 KO shows no later time amplification of VEGF (A) or IL-8 (B). NF-$\kappa\beta$ KO has an even more pronounced effect on both VEGF and IL-8 leading to reduction in both normal flux amounts as well as curtailing the signal amplification at later time points.

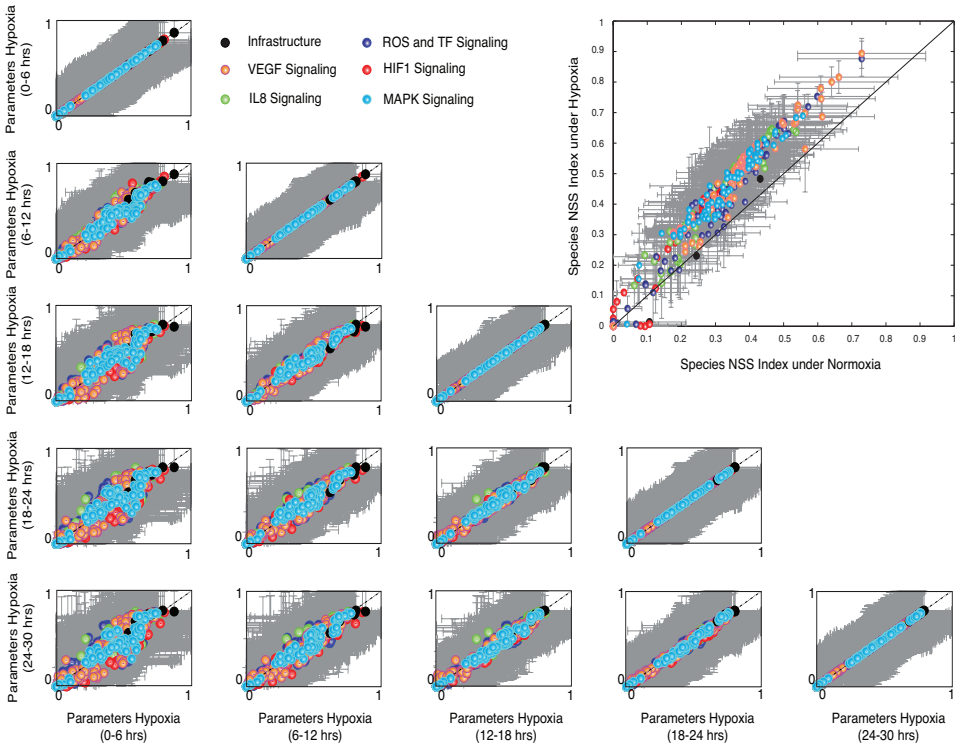


Figure A.3: Sensitivity plot for parameters with progression of time in EHR. Too large to include right here.

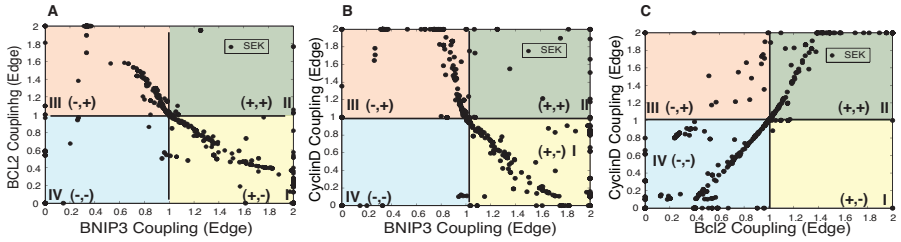


Figure A.4: Phenotypic behavior of EHR models upon Single Parameter Knockout: (A)-(C) Phenotypic behavior observed as a result of SPK. Phenotypic behavior was analyzed using downstream markers like BNIP3 (marker for cell death), BCL2 (marker for cell survival) and Cyclin D (marker for cell growth). 35% of SPK cases had a pro-death phenotype marked by increased (robustness coefficient > 1) marker levels of BNIP3 and decreased (robustness coefficient < 1) marker levels of BCL2. Around 20% of SPK cases had a pro-growth phenotype marked by increased (robustness coefficient > 1) marker levels of Cyclin D as compared to pro-death phenotype.

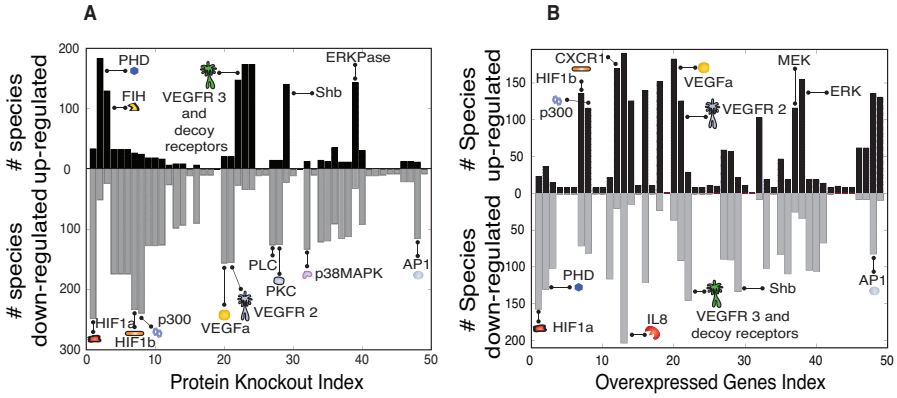


Figure A.5: Comparison of single gene knockout and gene over-expression in the EHR model: (A) Gene knockout studies revealed the critical players involved in up-regulation of the system as the removal of the HIF-1α degradation machinery, decoy receptors for the VEGF (VEGFR3) and the regulatory elements of the MAPK cascades (ERKPase). Early stage markers of the EHR signaling cascade like HIF-1α, HIF-1β, VEGF signaling initiators, PKC, p38MAPK, NF-$\kappa\beta$ and AP-1 were seen to have an enhanced down-regulatory effect. (B) Over-expression analysis results are complementary to that of knockout studies, revealing the enhanced role of HIF-1 signal initiation and regulation, VEGF decoy signaling and components of MAPK signaling. Increase in HIF-1β, p300 had a more pronounced effect in up-regulation.

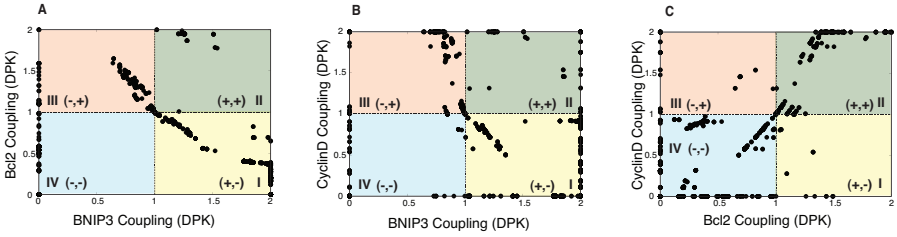


Figure A.6: Phenotypic behavior of EHR models upon Dual Gene Knockout: (A)-(C) Phenotypic behavior observed as a result of combinatorial gene knockouts. Sub-populations were identified that had a pro-survival (15%), pro-growth (15%) and pro-death (65%) phenotype. The positive correlation between pro-growth and pro-survival phenotypes could be attributed to the transcriptional role of AP1 and NF-$\kappa\beta$. The dominant phenotype of pro-death could be attributed to the HIF-1 dependent regulation of BNIP3 and NF-$\kappa\beta$ dependent regulation of BCL2.

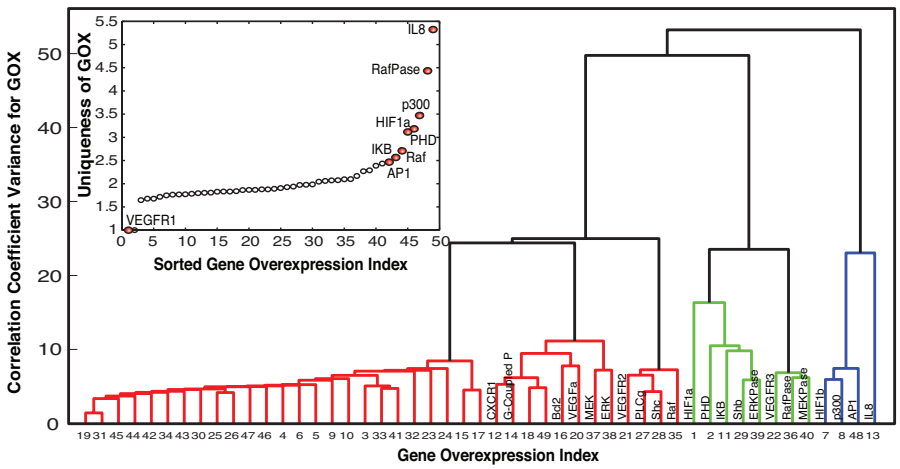


Figure A.7: Analysis of single gene overexpression study: We find that the IL-8 overexpression has the most pronounced effect, closely followed by regulators of AP-1, HIF-1 and NF-$\kappa\beta$ transcriptional activity.

BIBLIOGRAPHY

[1] M.D. Abramoff, P.J. Magelhaes, and S.J. Ram. Image processing with imagej. *Biophotonics Int*, 11:36–42, 2004.

[2] Jerry M Adams. Ways of dying: multiple pathways to apoptosis. *Genes Dev*, 17:2481–95, Oct 2003.

[3] Shaheen Ahmed and Ali Nawshad. Complexity in interpretation of embryonic epithelial-mesenchymal transition in response to transforming growth factor-beta signaling. *Cells Tissues Organs*, 185(1-3):131–45, 2007.

[4] Bree B Aldridge, John M Burke, Douglas A Lauffenburger, and Peter K Sorger. Physicochemical modelling of cell signalling pathways. *Nat Cell Biol*, 8(11):1195–1203, Nov 2006.

[5] A A Alizadeh, M B Eisen, R E Davis, C Ma, I S Lossos, A Rosenwald, J C Boldrick, H Sabet, T Tran, X Yu, J I Powell, L Yang, G E Marti, T Moore, J Hudson, Jr, L Lu, D B Lewis, R Tibshirani, G Sherlock, W C Chan, T C Greiner, D D Weisenburger, J O Armitage, R Warnke, R Levy, W Wilson, M R Grever, J C Byrd, D Botstein, P O Brown, and L M Staudt. Distinct types of diffuse large b-cell lymphoma identified by gene expression profiling. *Nature*, 403(6769):503–11, Feb 2000.

[6] E S Alnemri, D J Livingston, D W Nicholson, G Salvesen, N A Thornberry, W W Wong, and J Yuan. Human ice/ced-3 protease nomenclature. *Cell*, 87(2):171, Oct 1996.

[7] Uri Alon. Network motifs: theory and experimental approaches. *Nat Rev Genet*, 8(6):450–61, Jun 2007.

[8] O Alter, P O Brown, and D Botstein. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A*, 97(18):10101–6, Aug 2000.

[9] Orly Alter, Patrick O Brown, and David Botstein. Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms. *Proc Natl Acad Sci U S A*, 100(6):3351–6, Mar 2003.

[10] Kurosh Ameri and Adrian L Harris. Activating transcription factor 4. *Int J Biochem Cell Biol*, 40(1):14–21, 2008.

[11] Alexander R A Anderson and Vito Quaranta. Integrative mathematical oncology. *Nat Rev Cancer*, 8(3):227–34, Mar 2008.

[12] E B Anokhina and L B Buravkova. Mechanisms of regulation of transcription factor hif under hypoxia. *Biochemistry (Mosc)*, 75(2):151–8, Feb 2010.

[13] Eiichi Araki, Seiichi Oyadomari, and Masataka Mori. Endoplasmic reticulum stress and diabetes mellitus. *Intern Med*, 42(1):7–14, Jan 2003.

[14] Laura Arce, Kira T Pate, and Marian L Waterman. Groucho binds two conserved regions of lef-1 for hdac-dependent repression. *BMC Cancer*, 9:159, 2009.

[15] Richa Arya, Moushami Mallik, and Subhash C Lakhotia. Heat shock genes - integrating cell survival and death. *J Biosci*, 32(3):595–610, Apr 2007.

[16] Heike E Assmus, Ralf Herwig, Kwang-Hyun Cho, and Olaf Wolkenhauer. Dynamics of biological systems: role of systems biology in medical research. *Expert Rev Mol Diagn*, 6(6):891–902, Nov 2006.

[17] Chaitanya A Athale and Thomas S Deisboeck. The effects of egf-receptor density on multiscale tumor growth patterns. *J Theor Biol*, 238(4):771–9, Feb 2006.

[18] Roi Avraham and Yosef Yarden. Feedback regulation of egfr signalling: decision making by early and delayed loops. *Nat Rev Mol Cell Biol*, 12(2):104–17, Feb 2011.

[19] Hkan Axelson, Erik Fredlund, Marie Ovenberger, Göran Landberg, and Sven Phlman. Hypoxia-induced dedifferentiation of tumor cells–a mechanism behind heterogeneity and aggressiveness of solid tumors. *Semin Cell Dev Biol*, 16(4-5):554–63, 2005.

[20] Tomoya Baba, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A Datsenko, Masaru Tomita, Barry L Wanner, and Hirotada Mori. Construction of escherichia coli k-12 in-frame, single-gene knockout mutants: the keio collection. *Mol Syst Biol*, 2:2006.0008, 2006.

[21] Béatrice Bailly-Maitre, Constantino Fondevila, Fady Kaldas, Nathalie Droin, Fréderic Luciano, Jean-Ehrland Ricci, Rhonda Croxton, Maryla Krajewska, Juan M Zapata, Jerzy W Kupiec-Weglinski, Douglas Farmer,

and John C Reed. Cytoprotective gene bi-1 is required for intrinsic protection from endoplasmic reticulum stress and ischemia-reperfusion injury. *Proc Natl Acad Sci U S A*, 103(8):2809–14, Feb 2006.

[22] Yoshio Bando, Yoshitane Tsukamoto, Taichi Katayama, Kentaro Ozawa, Yasuko Kitao, Osamu Hori, David M Stern, Atsushi Yamauchi, and Satoshi Ogawa. Orp150/hsp12a protects renal tubular epithelium from ischemia-induced cell death. *FASEB J*, 18(12):1401–3, Sep 2004.

[23] A J Bannister, H J Brown, J A Sutherland, and T Kouzarides. Phosphorylation of the c-fos and c-jun hob1 motif stimulates its activation capacity. *Nucleic Acids Res*, 22(24):5173–6, Dec 1994.

[24] Ziv Bar-Joseph, Georg K Gerber, Tong Ihn Lee, Nicola J Rinaldi, Jane Y Yoo, François Robert, D Benjamin Gordon, Ernest Fraenkel, Tommi S Jaakkola, Richard A Young, and David K Gifford. Computational discovery of gene modules and regulatory networks. *Nat Biotechnol*, 21(11):1337–42, Nov 2003.

[25] Julia I Bárdos and Margaret Ashcroft. Negative and positive regulation of hif-1: a complex network. *Biochim Biophys Acta*, 1755(2):107–20, Jul 2005.

[26] D Battogtokh, D K Asch, M E Case, J Arnold, and H-B Schuttler. An ensemble method for identifying regulatory circuits with special reference to the qa gene cluster of neurospora crassa. *Proc Natl Acad Sci U S A*, 99(26):16904–16909, 2002.

[27] Véronique Baud and Michael Karin. Is nf-kappab a good target for cancer therapy? hopes and pitfalls. *Nat Rev Drug Discov*, 8(1):33–40, Jan 2009.

[28] DeannaLee M Beauvais, Brian J Ell, Andrea R McWhorter, and Alan C Rapraeger. Syndecan-1 regulates alphavbeta3 and alphavbeta5 integrin activation during angiogenesis and is blocked by synstatin, a novel peptide inhibitor. *J Exp Med*, 206(3):691–705, Mar 2009.

[29] Eric Becker, Lore Florin, Klaus Pfizenmaier, and Hitto Kaufmann. An xbp-1 dependent bottle-neck in production of igg subtype antibodies in chemically defined serum-free chinese hamster ovary (cho) fed-batch processes. *J Biotechnol*, 135(2):217–23, Jun 2008.

[30] A A Beg, W C Sha, R T Bronson, S Ghosh, and D Baltimore. Embryonic lethality and liver degeneration in mice lacking the rela component of nf-kappa b. *Nature*, 376(6536):167–70, Jul 1995.

[31] Steven L Bell and Bernhard Ø Palsson. Expa: a program for calculating extreme pathways in biochemical reaction networks. *Bioinformatics*, 21(8):1739–40, Apr 2005.

[32] Peter J Belmont, Archana Tadimalla, Wenqiong J Chen, Joshua J Martindale, Donna J Thuerauf, Marie Marcinko, Natalie Gude, Mark A Sussman, and Christopher C Glembotski. Coordination of growth and endoplasmic reticulum stress signaling by regulator of calcineurin 1 (rcan1), a novel atf6-inducible gene. *J Biol Chem*, 283(20):14012–21, May 2008.

[33] N F Bence, R M Sampat, and R R Kopito. Impairment of the ubiquitin-proteasome system by protein aggregation. *Science*, 292(5521):1552–5, May 2001.

[34] Arrigo De Benedetti and Jeremy R Graff. eIF-4E expression and its role in malignancies and metastases. *Oncogene*, 23(18):3189–3199, Apr 2004.

[35] Ina H Benoy, Roberto Salgado, Peter Van Dam, Katrien Geboers, Eric Van Marck, Simon Scharpé, Peter B Vermeulen, and Luc Y Dirix. Increased serum interleukin-8 in patients with early and metastatic breast cancer correlates with early dissemination and survival. *Clin Cancer Res*, 10(21):7157–62, Nov 2004.

[36] Sebastián Bernales, Kent L McDonald, and Peter Walter. Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response. *PLoS Biol*, 4(12):e423, Nov 2006.

[37] P Bernardi. Mitochondrial transport of cations: channels, exchangers, and permeability transition. *Physiol Rev*, 79(4):1127–55, Oct 1999.

[38] A Bertolotti, Y Zhang, L M Hendershot, H P Harding, and D Ron. Dynamic interaction of bip and er stress transducers in the unfolded-protein response. *Nat Cell Biol*, 2(6):326–32, Jun 2000.

[39] U S Bhalla and R Iyengar. Emergent properties of networks of biological signaling pathways. *Science*, 283(5400):381–7, Jan 1999.

[40] Meixia Bi, Christine Naczki, Marianne Koritzinsky, Diane Fels, Jaime Blais, Nianping Hu, Heather Harding, Isabelle Novoa, Mahesh Varia, James Raleigh, Donalyn Scheuner, Randal J Kaufman, John Bell, David Ron, Bradly G Wouters, and Constantinos Koumenis. Er stress-regulated translation increases tolerance to extreme hypoxia and promotes tumor growth. *EMBO J*, 24(19):3470–81, Oct 2005.

[41] Peter J Bickel and Aiyou Chen. A nonparametric view of network models and newman-girvan and other modularities. *Proc Natl Acad Sci U S A*, 106(50):21068–73, Dec 2009.

[42] Mary-Ann Bjornsti and Peter J Houghton. Lost in translation: dysregulation of cap-dependent translation and cancer. *Cancer Cell*, 5(6):519–523, Jun 2004.

[43] S Bläss, A Union, J Raymackers, F Schumann, U Ungethüm, S Müller-Steinbach, F De Keyser, J M Engel, and G R Burmester. The stress protein bip is overexpressed and is a major b and t cell target in rheumatoid arthritis. *Arthritis Rheum*, 44(4):761–71, Apr 2001.

[44] Michael L Blinov, James R Faeder, Byron Goldstein, and William S Hlavacek. Bionetgen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics*, 20(17):3289–91, Nov 2004.

[45] Nathan Blow. Systems biology: Untangling the protein web. *Nature*, 460(7253):415–8, Jul 2009.

[46] Nikolay Borisov, Edita Aksamitiene, Anatoly Kiyatkin, Stefan Legewie, Jan Berkhout, Thomas Maiwald, Nikolai P Kaimachnikov, Jens Timmer, Jan B Hoek, and Boris N Kholodenko. Systems-level interactions between insulin-EGF networks amplify mitogenic signaling. *Mol Syst Biol*, 5:256, 2009.

[47] Nicole Borth, Diethard Mattanovich, Renate Kunert, and Hermann Katinger. Effect of increased expression of protein disulfide isomerase and heavy chain binding protein on antibody secretion in a recombinant cho cell line. *Biotechnol Prog*, 21(1):106–11, 2005.

[48] Christopher D Bown, Jun-Feng Wang, Biao Chen, and L Trevor Young. Regulation of er stress proteins by valproate: therapeutic implications. *Bipolar Disord*, 4(2):145–51, Apr 2002.

[49] Michael Boyce, Kevin F Bryant, Céline Jousse, Kai Long, Heather P Harding, Donalyn Scheuner, Randal J Kaufman, Dawei Ma, Donald M Coen, David Ron, and Junying Yuan. A selective inhibitor of eif2alpha dephosphorylation protects cells from er stress. *Science*, 307(5711):935–9, Feb 2005.

[50] M Christiane Brahimi-Horn, Johanna Chiche, and Jacques Pouysségur. Hypoxia and cancer. *J Mol Med*, 85(12):1301–7, Dec 2007.

[51] Paul Brazhnik. Inferring gene networks from steady-state response to single-gene perturbations. *J Theor Biol*, 237(4):427–40, Dec 2005.

[52] Paul Brazhnik, Alberto de la Fuente, and Pedro Mendes. Gene networks: how to put the function in genomics. *Trends Biotechnol*, 20(11):467–72, Nov 2002.

[53] D M Brizel, S P Scully, J M Harrelson, L J Layfield, J M Bean, L R Prosnitz, and M W Dewhirst. Tumor oxygenation predicts for the likelihood of distant metastases in human soft tissue sarcoma. *Cancer Res*, 56(5):941–3, Mar 1996.

[54] Kevin S Brown and James P Sethna. Statistical mechanical approaches to models with many poorly known parameters. *Phys Rev E Stat Nonlin Soft Matter Phys*, 68(2 Pt 1):021904, 2003.

[55] Antonio Brú, Sonia Albertos, José Luis Subiza, José López García-Asenjo, and Isabel Brú. The universal dynamics of tumor growth. *Biophys J*, 85(5):2948–61, Nov 2003.

[56] Matthew H Brush, Douglas C Weiser, and Shirish Shenolikar. Growth arrest and dna damage-inducible protein gadd34 targets protein phosphatase 1 alpha to the endoplasmic reticulum and promotes dephosphorylation of the alpha subunit of eukaryotic translation initiation factor 2. *Mol Cell Biol*, 23:1292–303, Feb 2003.

[57] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, 10(3):186–98, Mar 2009.

[58] T R Burton and S B Gibson. The role of bcl-2 family member bnip3 in cell death and disease: Nipping at the heels of cell death. *Cell Death Differ*, 16(4):515–23, Apr 2009.

[59] Eugene C Butcher, Ellen L Berg, and Eric J Kunkel. Systems biology in drug discovery. *Nat Biotechnol*, 22(10):1253–9, Oct 2004.

[60] M. Calfon, H. Zeng, F. Urano, J.H. Till, S.R. Hubbard, H.P. Harding, S.G.

Clark, and D. Ron. IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature*, 415(6867):92–96, 2002.

[61] A Cano, M A Pérez-Moreno, I Rodrigo, A Locascio, M J Blanco, M G del Barrio, F Portillo, and M A Nieto. The transcription factor snail controls epithelial-mesenchymal transitions by repressing e-cadherin expression. *Nat Cell Biol*, 2(2):76–83, Feb 2000.

[62] Yang Cao, Daniel T Gillespie, and Linda R Petzold. The slow-scale stochastic simulation algorithm. *J Chem Phys*, 122(1):14116, Jan 2005.

[63] Julio J Caramelo, Olga A Castro, Gonzalo de Prat-Gay, and Armando J Parodi. The endoplasmic reticulum glucosyltransferase recognizes nearly native glycoprotein folding intermediates. *J Biol Chem*, 279(44):46280–5, Oct 2004.

[64] Jennifer S Carew, Steffan T Nawrocki, Yelena V Krupnik, Kenneth Dunner, Jr, David J McConkey, Michael J Keating, and Peng Huang. Targeting endoplasmic reticulum protein transport: a novel strategy to kill malignant b cells and overcome fludarabine resistance in cll. *Blood*, 107(1):222–31, Jan 2006.

[65] J M Carlson and John Doyle. Complexity and robustness. *Proc Natl Acad Sci U S A*, 99 Suppl 1:2538–45, Feb 2002.

[66] P Carmeliet and R K Jain. Angiogenesis in cancer and other diseases. *Nature*, 407(6801):249–57, Sep 2000.

[67] Peter Carmeliet and Marc Tessier-Lavigne. Common mechanisms of nerve and blood vessel wiring. *Nature*, 436(7048):193–200, Jul 2005.

[68] Etienne Caron, Samik Ghosh, Yukiko Matsuoka, Dariel Ashton-Beaucage, Marc Therrien, Sébastien Lemieux, Claude Perreault, Philippe P Roux, and Hiroaki Kitano. A comprehensive map of the mTOR signaling network. *Mol Syst Biol*, 6:453, Dec 2010.

[69] Oriol Casanovas, Daniel J Hicklin, Gabriele Bergers, and Douglas Hanahan. Drug resistance by evasion of antiangiogenic targeting of vegf signaling in late-stage pancreatic islet tumors. *Cancer Cell*, 8(4):299–309, Oct 2005.

[70] Anirikh Chakrabarti, Joshua P Lequieu, Alvin N Kumarga, Brian Manzi,

Maja Stojiljkovic, Monroe Chen, Nicholas B Guros, Robert A Dromms, Ryan Tasseff, Stroock Abraham Verbridge, Scott S, Claudia Fischbach, and Jeffrey D Varner. Analysis of the eukaryotic hypoxic response revealed hif-1, ap-1 and nf-kb mediated crosstalk between il-8 and vegf signaling cascades. 2011.

[71] Claudine Chaouiya. Petri net modelling of biological networks. *Brief Bioinform*, 8(4):210–9, Jul 2007.

[72] Dina Chelouche-Lev, Claudia P Miller, Carmen Tellez, Maribelis Ruiz, Menashe Bar-Eli, and Janet E Price. Different signalling pathways regulate vegf and il-8 expression in breast cancer: implications for therapy. *Eur J Cancer*, 40(16):2509–18, Nov 2004.

[73] Dan Chen and Qiang Zhou. Caspase cleavage of bimel triggers a positive feedback amplification of apoptotic signaling. *Proc Natl Acad Sci U S A*, 101(5):1235–40, Feb 2004.

[74] William W Chen, Birgit Schoeberl, Paul J Jasper, Mario Niepel, Ulrik B Nielsen, Douglas A Lauffenburger, and Peter K Sorger. Input-output behavior of erbb signaling pathways as revealed by a mass action model trained against dynamic data. *Mol Syst Biol*, 5:239, 2009.

[75] Xi Chen, Jingshi Shen, and Ron Prywes. The luminal domain of atf6 senses endoplasmic reticulum (er) stress and causes translocation of atf6 from the er to the golgi. *J Biol Chem*, 277(15):13045–52, Apr 2002.

[76] Seung-Wook Chung, Fayth L Miles, Robert A Sikes, Carlton R Cooper, Mary C Farach-Carson, and Babatunde A Ogunnaike. Quantitative modeling and analysis of the transforming growth factor beta signaling pathway. *Biophys J*, 96(5):1733–50, Mar 2009.

[77] Sven Cichon, Silvia Buervenich, George Kirov, Nirmala Akula, Albena Dimitrova, Elaine Green, Johannes Schumacher, Norman Klopp, Tim Becker, Stephanie Ohlraun, Thomas G Schulze, Monja Tullius, Magdalena M Gross, Lisa Jones, Stefan Krastev, Ivan Nikolov, Marian Hamshere, Ian Jones, Piotr M Czerski, Anna Leszczynska-Rodziewicz, Pawel Kapelski, Ann Van Den Bogaert, Thomas Illig, Joanna Hauser, Wolfgang Maier, Wade Berrettini, William Byerley, William Coryell, Elliot S Gershon, John R Kelsoe, Melvin G McInnis, Dennis L Murphy, John I Nurnberger, Theodore Reich, William Scheftner, Michael C O'Donovan, Peter Propping, Michael J Owen, Marcella Rietschel, Markus M Nöthen, Francis J McMahon, and Nick Craddock. Lack of support for a genetic

association of the xbp1 promoter polymorphism with bipolar disorder in probands of european origin. *Nat Genet*, 36(8):783–4; author reply 784–5, Aug 2004.

[78] P Codogno and A J Meijer. Autophagy and signaling: their role in cell survival and cell death. *Cell Death Differ*, 12 Suppl 2:1509–18, Nov 2005.

[79] Joshua Colvin, Michael I Monine, Ryan N Gutenkunst, William S Hlavacek, Daniel D Von Hoff, and Richard G Posner. Rulemonkey: software for stochastic simulation of rule-based models. *BMC Bioinformatics*, 11:404, 2010.

[80] H Conzelmann, J Saez-Rodriguez, T Sauter, E Bullinger, F Allgöwer, and E D Gilles. Reduction of mathematical models of signal transduction networks: simulation-based approach applied to egf receptor signalling. *Syst Biol (Stevenage)*, 1(1):159–69, Jun 2004.

[81] Holger Conzelmann, Dirk Fey, and Ernst D Gilles. Exact model reduction of combinatorial reaction networks. *BMC Syst Biol*, 2:78, 2008.

[82] Nadia Corazza, Sabine Jakob, Corinne Schaer, Steffen Frese, Adrian Keogh, Deborah Stroka, Daniela Kassahn, Ralph Torgler, Christoph Mueller, Pascal Schneider, and Thomas Brunner. Trail receptor-mediated jnk activation and bim phosphorylation critically regulate fas-mediated liver damage and lethality. *J Clin Invest*, 116(9):2493–9, Sep 2006.

[83] V M Corrigall, M D Bodman-Smith, M S Fife, B Canas, L K Myers, P Wooley, C Soh, N A Staines, D J Pappin, S E Berlo, W van Eden, R van Der Zee, J S Lanchbury, and G S Panayi. The human endoplasmic reticulum molecular chaperone bip is an autoantigen for rheumatoid arthritis and prevents the induction of experimental arthritis. *J Immunol*, 166(3):1492–8, Feb 2001.

[84] Sonia Cortassa, Miguel A Aon, Brian O'Rourke, Robert Jacques, Hsiang-Jer Tseng, Eduardo Marbán, and Raimond L Winslow. A computational model integrating electrophysiology, contraction, and mitochondrial bioenergetics in the ventricular myocyte. *Biophys J*, 91(4):1564–89, Aug 2006.

[85] J S Cox, C E Shamu, and P Walter. Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell*, 73(6):1197–206, Jun 1993.

[86] Carolyn Culver, Anders Sundqvist, Sharon Mudie, Andrew Melvin, Dimitris Xirodimas, and Sonia Rocha. Mechanism of hypoxia-induced nf-kappab. *Mol Cell Biol*, 30(20):4901–21, Oct 2010.

[87] Michal Dabrowski, Norbert Dojer, Malgorzata Zawadzka, Jakub Mieczkowski, and Bozena Kaminska. Comparative analysis of cis-regulation following stroke and seizures in subspaces of conserved eigensystems. *BMC Syst Biol*, 4:86, 2010.

[88] Cecilia D'Alessio, Julio J Caramelo, and Armando J Parodi. Udp-glc:glycoprotein glucosyltransferase-glucosidase ii, the ying-yang of the er quality control. *Semin Cell Dev Biol*, 21(5):491–9, Jul 2010.

[89] Stephen G Dann, Anand Selvaraj, and George Thomas. mTOR Complex1-S6K1 signaling: at the crossroads of obesity, diabetes and cancer. *Trends Mol Med*, 13(6):252–259, Jun 2007.

[90] Madhukar S Dasika, Anthony Burgard, and Costas D Maranas. A computational framework for the topological analysis and targeted disruption of signal transduction networks. *Biophys J*, 91(1):382–98, Jul 2006.

[91] R Davis, K Schooley, B Rasmussen, J Thomas, and P Reddy. Effect of pdi overexpression on recombinant protein secretion in cho cells. *Biotechnol Prog*, 16(5):736–43, 2000.

[92] Ted M Dawson and Valina L Dawson. Rare genetic mutations shed light on the pathogenesis of parkinson disease. *J Clin Invest*, 111(2):145–51, Jan 2003.

[93] Sarmishtha De, Olga Razorenova, Noel Patrick McCabe, Timothy O'Toole, Jun Qin, and Tatiana V Byzova. Vegf-integrin interplay controls tumor growth and vascularization. *Proc Natl Acad Sci U S A*, 102(21):7589–94, May 2005.

[94] Giovanna De Chiara, Maria Elena Marcocci, Maria Torcia, Maria Lucibello, Paolo Rosini, Paolo Bonini, Yukiro Higashimoto, Gianluca Damonte, Andrea Armirotti, Sarah Amodei, Anna Teresa Palamara, Tommaso Russo, Enrico Garaci, and Federico Cozzolino. Bcl-2 phosphorylation by p38 mapk: identification of target sites and biologic consequences. *J Biol Chem*, 281(30):21353–61, Jul 2006.

[95] Hidde de Jong. Modeling and simulation of genetic regulatory systems: a literature review. *J Comput Biol*, 9(1):67–103, 2002.

[96] Alberto de la Fuente, Paul Brazhnik, and Pedro Mendes. Linking the genes: inferring quantitative gene networks from microarray data. *Trends Genet*, 18(8):395–8, Aug 2002.

[97] Antonella De Luca, Adele Carotenuto, Annamaria Rachiglio, Marianna Gallo, Monica R Maiello, Donatella Aldinucci, Antonio Pinto, and Nicola Normanno. The role of the egfr signaling in tumor microenvironment. *J Cell Physiol*, 214(3):559–67, Mar 2008.

[98] Antonella De Luca, Marianna Gallo, Donatella Aldinucci, Domenico Ribatti, Luana Lamura, Amelia D'Alessio, Rosaria De Filippi, Antonio Pinto, and Nicola Normanno. Role of the egfr ligand/receptor system in the secretion of angiogenic factors in mesenchymal stem cells. *J Cell Physiol*, 226(8):2131–8, Aug 2011.

[99] D Decaudin, S Geley, T Hirsch, M Castedo, P Marchetti, A Macho, R Kofler, and G Kroemer. Bcl-2 and bcl-xl antagonize the mitochondrial dysfunction preceding nuclear apoptosis induced by chemotherapeutic agents. *Cancer Res*, 57(1):62–7, Jan 1997.

[100] Thomas S Deisboeck. Personalizing medicine: a systems biology perspective. *Mol Syst Biol*, 5:249, 2009.

[101] Nicolas Demaurex and Clark Distelhorst. Cell biology. apoptosis–the calcium connection. *Science*, 300(5616):65–7, Apr 2003.

[102] B Dérijard, J Raingeaud, T Barrett, I H Wu, J Han, R J Ulevitch, and R J Davis. Independent human map-kinase signal transduction pathways defined by mek and mkk isoforms. *Science*, 267(5198):682–5, Feb 1995.

[103] Rik Derynck and Ying E Zhang. Smad-dependent and smad-independent pathways in tgf-beta family signalling. *Nature*, 425(6958):577–84, Oct 2003.

[104] T G Dewey and D J Galas. Dynamic models of gene expression and classification. *Funct Integr Genomics*, 1(4):269–78, Mar 2001.

[105] R.P. Dickinson and R.J. Gelinas. Sensitivity analysis of ordinary differential equation systems—a direct method. *J. Comput. Phys*, 21(2), 1976.

[106] R L Dillon, D E White, and W J Muller. The phosphatidyl inositol 3-kinase signaling network: implications for human breast cancer. *Oncogene*, 26(9):1338–45, Feb 2007.

[107] Richard J Dimelow and Stephen J Wilkinson. Control of translation initiation: a model-based analysis from limited experimental data. *J R Soc Interface*, 6(30):51–61, Jan 2009.

[108] Clark W Distelhorst and Gordon C Shore. Bcl-2 and calcium: controversy beneath the surface. *Oncogene*, 23(16):2875–80, Apr 2004.

[109] J Dufraine, Y Funahashi, and J Kitajewski. Notch signaling regulates tumor angiogenesis by diverse mechanisms. *Oncogene*, 27(38):5132–7, Sep 2008.

[110] R E Durand. Multicell spheroids as a model for cell kinetic studies. *Cell Tissue Kinet*, 23(3):141–59, May 1990.

[111] John B Easton, Raushan T Kurmasheva, and Peter J Houghton. IRS-1: auditing the effectiveness of mTOR inhibitors. *Cancer Cell*, 9(3):153–155, Mar 2006.

[112] Thomas Eissing, Lars Kuepfer, Corina Becker, Michael Block, Katrin Coboeken, Thomas Gaub, Linus Goerlitz, Juergen Jaeger, Roland Loosen, Bernd Ludewig, Michaela Meyer, Christoph Niederalt, Michael Sevestre, Hans-Ulrich Siegmund, Juri Solodenko, Kirstin Thelen, Ulrich Telle, Wolfgang Weiss, Thomas Wendl, Stefan Willmann, and Joerg Lippert. A computational systems biology software platform for multiscale modeling and simulation: integrating whole-body physiology, disease biology, and molecular reaction networks. *Front Physiol*, 2:4, 2011.

[113] M. Elchebly, P. Payette, E. Michaliszyn, W. Cromlish, S. Collins, A. L. Loy, D. Normandin, A. Cheng, J. Himms-Hagen, C. C. Chan, C. Ramachandran, M. J. Gresser, M. L. Tremblay, and B. P. Kennedy. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science*, 283(5407):1544–1548, Mar 1999.

[114] L Ellgaard, M Molinari, and A Helenius. Setting the standards: quality control in the secretory pathway. *Science*, 286(5446):1882–8, Dec 1999.

[115] L Ellgaard, R Riek, D Braun, T Herrmann, A Helenius, and K Wüthrich. Three-dimensional structure topology of the calreticulin p-domain based on nmr assignment. *FEBS Lett*, 488(1-2):69–73, Jan 2001.

[116] L Ellgaard, R Riek, T Herrmann, P Güntert, D Braun, A Helenius, and K Wüthrich. Nmr structure of the calreticulin p-domain. *Proc Natl Acad Sci U S A*, 98(6):3133–8, Mar 2001.

[117] Lars Ellgaard and Ari Helenius. Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol*, 4(3):181–91, Mar 2003.

[118] Jeffrey A Engelman. Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nat Rev Cancer*, 9(8):550–562, Aug 2009.

[119] Adam J Engler, Patrick O Humbert, Bernhard Wehrle-Haller, and Valerie M Weaver. Multiscale modeling of form and function. *Science*, 324(5924):208–12, Apr 2009.

[120] J R Faeder, M L Blinov, B Goldstein, and W S Hlavacek. Combinatorial complexity and dynamical restriction of network flows in signal transduction. *Syst Biol (Stevenage)*, 2(1):5–15, Mar 2005.

[121] Iman Famili and Bernhard O Palsson. Systemic metabolic reactions are obtained by singular value decomposition of genome-scale stoichiometric matrices. *J Theor Biol*, 224(1):87–96, Sep 2003.

[122] Dana Faratian, Alexey Goltsov, Galina Lebedeva, Anatoly Sorokin, Stuart Moodie, Peter Mullen, Charlene Kay, In Hwa Um, Simon Langdon, Igor Goryanin, and David J Harrison. Systems biology reveals new strategies for personalizing cancer medicine and confirms the role of PTEN in resistance to trastuzumab. *Cancer Res*, 69(16):6713–6720, Aug 2009.

[123] Amjad Farooq and Ming-Ming Zhou. Structure and regulation of mapk phosphatases. *Cell Signal*, 16(7):769–79, Jul 2004.

[124] A N Fedorov and T O Baldwin. Cotranslational protein folding. *J Biol Chem*, 272(52):32715–8, Dec 1997.

[125] Napoleone Ferrara. Role of vascular endothelial growth factor in physiologic and pathologic angiogenesis: therapeutic implications. *Semin Oncol*, 29(6 Suppl 16):10–4, Dec 2002.

[126] Napoleone Ferrara. Vegf as a therapeutic target in cancer. *Oncology*, 69 Suppl 3:11–6, 2005.

[127] Napoleone Ferrara, Kenneth J Hillan, Hans-Peter Gerber, and William Novotny. Discovery and development of bevacizumab, an anti-vegf antibody for treating cancer. *Nat Rev Drug Discov*, 3(5):391–400, May 2004.

[128] Diane C Fingar and John Blenis. Target of rapamycin (TOR): an integrator

of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression. *Oncogene*, 23(18):3151–3171, Apr 2004.

[129] Claudia Fischbach, Ruth Chen, Takuya Matsumoto, Tobias Schmelzle, Joan S Brugge, Peter J Polverini, and David J Mooney. Engineering tumors with 3d scaffolds. *Nat Methods*, 4(10):855–60, Oct 2007.

[130] Claudia Fischbach, Hyun Joon Kong, Susan X Hsiong, Marta B Evangelista, Will Yuen, and David J Mooney. Cancer cell angiogenic capability is regulated by 3d culture and integrin engagement. *Proc Natl Acad Sci U S A*, 106(2):399–404, Jan 2009.

[131] L Flohé, R Brigelius-Flohé, C Saliou, M G Traber, and L Packer. Redox regulation of nf-kappa b activation. *Free Radic Biol Med*, 22(6):1115–26, 1997.

[132] Bruno D Fonseca, Ewan M Smith, Vivian H-Y Lee, Carol MacKintosh, and Christopher G Proud. PRAS40 is a target for mammalian target of rapamycin complex 1 and is required for signaling downstream of this complex. *J Biol Chem*, 282(34):24514–24524, Aug 2007.

[133] C.M. Fonseca, P.J. Fleming, et al. Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In *Proceedings of the fifth international conference on genetic algorithms*, volume 423. Citeseer, 1993.

[134] Sonya G Fonseca, Mark Burcin, Jesper Gromada, and Fumihiko Urano. Endoplasmic reticulum stress in beta-cells and development of diabetes. *Curr Opin Pharmacol*, 9(6):763–70, Dec 2009.

[135] Sonya G Fonseca, Mariko Fukuma, Kathryn L Lipson, Linh X Nguyen, Jenny R Allen, Yoshitomo Oka, and Fumihiko Urano. Wfs1 is a novel component of the unfolded protein response and maintains homeostasis of the endoplasmic reticulum in pancreatic beta-cells. *J Biol Chem*, 280(47):39609–15, Nov 2005.

[136] Christian V Forst. Network genomics–a novel approach for the analysis of biological systems in the post-genomic era. *Mol Biol Rep*, 29(3):265–80, Sep 2002.

[137] Jochen Förster, Iman Famili, Patrick Fu, Bernhard Ø Palsson, and Jens Nielsen. Genome-scale reconstruction of the saccharomyces cerevisiae metabolic network. *Genome Res*, 13(2):244–53, Feb 2003.

[138] A M Fra, C Fagioli, D Finazzi, R Sitia, and C M Alberini. Quality control of er synthesized proteins: an exposed thiol group as a three-way switch mediating assembly, retention and degradation. *EMBO J*, 12(12):4755–61, Dec 1993.

[139] Marcela Franco, Pernilla Roswall, Eliane Cortez, Douglas Hanahan, and Kristian Pietras. Pericytes promote endothelial cell survival through induction of autocrine vegf-a signaling and bcl-w expression. *Blood*, 118(10):2906–17, Sep 2011.

[140] Ariane Freund, Corine Chauveau, Jean-Paul Brouillet, Annick Lucas, Matthieu Lacroix, Anne Licznar, Françoise Vignon, and Gwendal Lazennec. Il-8 expression and its possible relationship with estrogen-receptor-negative status of breast cancer cells. *Oncogene*, 22(2):256–65, Jan 2003.

[141] Alain Friboulet and Daniel Thomas. Systems biology-an interdisciplinary approach. *Biosens Bioelectron*, 20(12):2404–7, Jun 2005.

[142] Eva-Maria Frickel, Roland Riek, Ilian Jelesarov, Ari Helenius, Kurt Wuthrich, and Lars Ellgaard. Trosy-nmr reveals interaction between erp57 and the tip of the calreticulin p-domain. *Proc Natl Acad Sci U S A*, 99(4):1954–9, Feb 2002.

[143] Hermann B Frieboes, Mary E Edgerton, John P Fruehauf, Felicity R A J Rose, Lisa K Worrall, Robert A Gatenby, Mauro Ferrari, and Vittorio Cristini. Prediction of drug response in breast cancer using integrative experimental/computational modeling. *Cancer Res*, 69(10):4484–92, May 2009.

[144] Yong Fu, Jianze Li, and Amy S Lee. Grp78/bip inhibits endoplasmic reticulum bik and protects human breast cancer cells against estrogen starvation-induced apoptosis. *Cancer Res*, 67(8):3734–40, Apr 2007.

[145] Takashi Fujimoto, Masamitsu Onda, Hisaki Nagai, Takemitsu Nagahata, Kenji Ogawa, and Mitsuru Emi. Upregulation and overexpression of human x-box binding protein 1 (hxbp-1) gene in primary breast cancers. *Breast Cancer*, 10(4):301–6, 2003.

[146] Eriko Fujita, Yoriko Kouroku, Atsushi Isoai, Hiromichi Kumagai, Akifumi Misutani, Chie Matsuda, Yukiko K Hayashi, and Takashi Momoi. Two endoplasmic reticulum-associated degradation (erad) systems for the novel variant of the mutant dysferlin: ubiquitin/proteasome erad(i) and autophagy/lysosome erad(ii). *Hum Mol Genet*, 16(6):618–29, Mar 2007.

[147] M Fussenegger, J E Bailey, and J Varner. A mathematical model of caspase function in apoptosis. *Nat Biotechnol*, 18(7):768–74, Jul 2000.

[148] K. G. Gadkar, J Varner, and F J Doyle. Model identification of signal transduction networks from data using a state regulator problem. *Syst Biol (Stevenage)*, 2(1):17–30, Mar 2005.

[149] K.G. Gadkar, F.J. Doyle III, T.J. Crowley, and J.D. Varner. Cybernetic model predictive control of a continuous bioreactor with cell recycle. *Biotechnology progress*, 19(5):1487–1497, 2003.

[150] Boyi Gan, Zara K Melkoumian, Xiaoyang Wu, Kun-Liang Guan, and Jun-Lin Guan. Identification of fip200 interaction with the tsc1-tsc2 complex and its role in regulation of cell size control. *J Cell Biol*, 170(3):379–389, Aug 2005.

[151] Attila Garami, Fried J T Zwartkruis, Takahiro Nobukuni, Manel Joaquin, Marta Roccio, Hugo Stocker, Sara C Kozma, Ernst Hafen, Johannes L Bos, and George Thomas. Insulin activation of Rheb, a mediator of mTOR/S6K/4E-BP signaling, is inhibited by TSC1 and 2. *Mol Cell*, 11(6):1457–1466, Jun 2003.

[152] Peter S Gargalovic, Nima M Gharavi, Michael J Clark, Joanne Pagnon, Wen-Pin Yang, Aiqing He, Amy Truong, Tamar Baruch-Oren, Judith A Berliner, Todd G Kirchgessner, and Aldons J Lusis. The unfolded protein response is an important regulator of inflammatory genes in endothelial cells. *Arterioscler Thromb Vasc Biol*, 26(11):2490–6, Nov 2006.

[153] Barbara Garmy-Susini and Judith A Varner. Roles of integrins in tumor angiogenesis and lymphangiogenesis. *Lymphat Res Biol*, 6(3-4):155–63, 2008.

[154] R A Gatenby and E T Gawlinski. A reaction-diffusion model of cancer invasion. *Cancer Res*, 56(24):5745–53, Dec 1996.

[155] Catherine E Gatza, Sun Young Oh, and Gerard C Blobe. Roles for the type iii tgf-beta receptor in human cancer. *Cell Signal*, 22(8):1163–74, Aug 2010.

[156] Fátima Gebauer and Matthias W Hentze. Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol*, 5(10):827–35, Oct 2004.

[157] MJ Gething. Role and regulation of the ER chaperone BiP. *Semin Cell Dev Biol*, 10:465–472, 1999.

[158] David Gilbert, Hendrik Fuss, Xu Gu, Richard Orton, Steve Robinson, Vladislav Vyshemirsky, Mary Jo Kurth, C Stephen Downes, and Werner Dubitzky. Computational methodologies for modelling, analysis and simulation of signalling networks. *Brief Bioinform*, 7(4):339–53, Dec 2006.

[159] Daniel T Gillespie. Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem*, 58:35–55, 2007.

[160] Joell J Gills, Jaclyn Lopiccolo, Junji Tsurutani, Robert H Shoemaker, Carolyn J M Best, Mones S Abu-Asab, Jennifer Borojerdi, Noel A Warfel, Erin R Gardner, Matthew Danish, M Christine Hollander, Shigeru Kawabata, Maria Tsokos, William D Figg, Patricia S Steeg, and Phillip A Dennis. Nelfinavir, a lead hiv protease inhibitor, is a broad-spectrum, anticancer agent that induces endoplasmic reticulum stress, autophagy, and apoptosis in vitro and in vivo. *Clin Cancer Res*, 13(17):5183–94, Sep 2007.

[161] A. C. Gingras, S. G. Kennedy, M. A. O'Leary, N. Sonenberg, and N. Hay. 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. *Genes Dev*, 12(4):502–513, Feb 1998.

[162] A. C. Gingras, B. Raught, S. P. Gygi, A. Niedzwiecka, M. Miron, S. K. Burley, R. D. Polakiewicz, A. Wyslouch-Cieszynska, R. Aebersold, and N. Sonenberg. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. *Genes Dev*, 15(21):2852–2864, Nov 2001.

[163] M Girvan and M E J Newman. Community structure in social and biological networks. *Proc Natl Acad Sci U S A*, 99(12):7821–6, Jun 2002.

[164] Christopher C Glembotski. Endoplasmic reticulum stress in the heart. *Circ Res*, 101(10):975–84, Nov 2007.

[165] Christopher C Glembotski. The role of the unfolded protein response in the heart. *J Mol Cell Cardiol*, 44:453–9, Mar 2008.

[166] T R Golub, D K Slonim, P Tamayo, C Huard, M Gaasenbeek, J P Mesirov, H Coller, M L Loh, J R Downing, M A Caligiuri, C D Bloomfield, and E S Lander. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, 286(5439):531–7, Oct 1999.

[167] T Gotoh, K Terada, and M Mori. hsp70-dnaj chaperone pairs prevent nitric oxide-mediated apoptosis in raw 264.7 macrophages. *Cell Death Differ*, 8(4):357–66, Apr 2001.

[168] Y Gotoh and J A Cooper. Reactive oxygen species- and dimerization-induced activation of apoptosis signal-regulating kinase 1 in tumor necrosis factor-alpha signal transduction. *J Biol Chem*, 273(28):17477–82, Jul 1998.

[169] Sara Gracie, Craig Pennell, Gunvor Ekman-Orderberg, Stephen Lye, James McManaman, Scott Williams, Lyle Palmer, Maureen Kelley, Ram Menon, Michael Gravett, and The Prebic '-Omics' Research Group. An integrated systems biology approach to the study of preterm birth using "-omic" technology - a guideline for research. *BMC Pregnancy Childbirth*, 11(1):71, Oct 2011.

[170] C J Green, P Lichtlen, N T Huynh, M Yanovsky, K R Laderoute, W Schaffner, and B J Murphy. Placenta growth factor gene expression is induced by hypoxia in fibroblasts: a central role for metal transcription factor-1. *Cancer Res*, 61(6):2696–703, Mar 2001.

[171] Feng Gu, Duc Thang Nguyên, Matthew Stuible, Nadia Dubé, Michel L Tremblay, and Eric Chevet. Protein-tyrosine phosphatase 1b potentiates ire1 signaling during endoplasmic reticulum stress. *J Biol Chem*, 279(48):49689–93, Nov 2004.

[172] J J Gu, Z Wang, R Reeves, and N S Magnuson. Pim1 phosphorylates and negatively regulates ask1-mediated apoptosis. *Oncogene*, 28(48):4261–71, Dec 2009.

[173] Philippe Gual, Yannick Le Marchand-Brustel, and Jean-François Tanti. Positive and negative regulation of insulin signaling through IRS-1 phosphorylation. *Biochimie*, 87(1):99–109, Jan 2005.

[174] Roger Guimerà and Marta Sales-Pardo. Missing and spurious interactions and the reconstruction of complex networks. *Proc Natl Acad Sci U S A*, 106(52):22073–8, Dec 2009.

[175] Sanjeev Gupta, Ayswaria Deepti, Shane Deegan, Fernanda Lisbona, Claudio Hetz, and Afshin Samali. Hsp72 protects cells from er stress-induced apoptosis via enhancement of ire1alpha-xbp1 signaling through a physical interaction. *PLoS Biol*, 8(7):e1000410, 2010.

[176] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol*, 3(10):1871–78, Oct 2007.

[177] Emilie Vander Haar, Seong-Il Lee, Sricharan Bandhakavi, Timothy J Griffin, and Do-Hyung Kim. Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40. *Nat Cell Biol*, 9(3):316–323, Mar 2007.

[178] T W Hai, F Liu, W J Coukos, and M R Green. Transcription factor atf cdna clones: an extensive family of leucine zipper proteins able to selectively form dna-binding heterodimers. *Genes Dev*, 3(12B):2083–90, Dec 1989.

[179] R B Hamanaka, E Bobrovnikova-Marjon, X Ji, S A Liebhaber, and J A Diehl. Perk-dependent regulation of iap translation during er stress. *Oncogene*, 28(6):910–20, Feb 2009.

[180] C Hammond, I Braakman, and A Helenius. Role of n-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc Natl Acad Sci U S A*, 91(3):913–7, Feb 1994.

[181] D Hanahan and R A Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, Jan 2000.

[182] Julia Handl, Douglas B Kell, and Joshua Knowles. Multiobjective optimization in bioinformatics and computational biology. *IEEE/ACM Trans Comput Biol Bioinform*, 4(2):279–92, 2007.

[183] H P Harding, H Zeng, Y Zhang, R Jungries, P Chung, H Plesken, D D Sabatini, and D Ron. Diabetes mellitus and exocrine pancreatic dysfunction in perk-/- mice reveals a role for translational control in secretory cell survival. *Mol Cell*, 7(6):1153–63, Jun 2001.

[184] H P Harding, Y Zhang, and D Ron. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature*, 397(6716):271–4, Jan 1999.

[185] Heather P Harding, Yuhong Zhang, Huiquing Zeng, Isabel Novoa, Phoebe D Lu, Marcella Calfon, Navid Sadri, Chi Yun, Brian Popko, Richard Paules, David F Stojdl, John C Bell, Thore Hettmann, Jeffrey M Leiden, and David Ron. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol Cell*, 11(3):619–33, Mar 2003.

[186] Simon Hardy and Pierre N Robillard. Petri net-based method for the analysis of the dynamics of signal propagation in signaling pathways. *Bioinformatics*, 24(2):209–17, Jan 2008.

[187] Judith H Harmey and David Bouchier-Hayes. Vascular endothelial growth factor (vegf), a survival factor for tumour cells: implications for anti-angiogenic therapy. *Bioessays*, 24(3):280–3, Mar 2002.

[188] Adrian L Harris. Hypoxia–a key regulatory factor in tumour growth. *Nat Rev Cancer*, 2(1):38–47, Jan 2002.

[189] Nissim Hay and Nahum Sonenberg. Upstream and downstream of mTOR. *Genes Dev*, 18(16):1926–1945, Aug 2004.

[190] K Haze, T Okada, H Yoshida, H Yanagi, T Yura, M Negishi, and K Mori. Identification of the g13 (camp-response-element-binding protein-related protein) gene product related to activating transcription factor 6 as a transcriptional activator of the mammalian unfolded protein response. *Biochem J*, 355(Pt 1):19–28, Apr 2001.

[191] K Haze, H Yoshida, H Yanagi, T Yura, and K Mori. Mammalian transcription factor atf6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol Biol Cell*, 10:3787–99, Nov 1999.

[192] D N Hebert, B Foellmer, and A Helenius. Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell*, 81(3):425–33, May 1995.

[193] D N Hebert, B Foellmer, and A Helenius. Calnexin and calreticulin promote folding, delay oligomerization and suppress degradation of influenza hemagglutinin in microsomes. *EMBO J*, 15(12):2961–8, Jun 1996.

[194] Daniel N Hebert, Riccardo Bernasconi, and Maurizio Molinari. Erad substrates: which way out? *Semin Cell Dev Biol*, 21(5):526–32, Jul 2010.

[195] Daniel N Hebert, Scott C Garman, and Maurizio Molinari. The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. *Trends Cell Biol*, 15(7):364–70, Jul 2005.

[196] Monika Heiner and K Sriram. Structural analysis to determine the core of hypoxia response network. *PLoS One*, 5(1):e8600, 2010.

[197] Laura M Heiser, Nicholas J Wang, Carolyn L Talcott, Keith R Laderoute, Merrill Knapp, Yinghui Guan, Zhi Hu, Safiyyah Ziyad, Barbara L Weber, Sylvie Laquerre, Jeffrey R Jackson, Richard F Wooster, Wen Lin Kuo, Joe W Gray, and Paul T Spellman. Integrated analysis of breast cancer cell lines reveals unique signaling pathways. *Genome Biol*, 10(3):R31, 2009.

[198] A. Helenius, ES Trombetta, DN Hebert, and JF Simons. Calnexin, calreticulin and the folding of glycoproteins. *Trends Cell Biol.*, 7(5):193–200, 1997.

[199] Ari Helenius and Markus Aebi. Roles of n-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem*, 73:1019–49, 2004.

[200] R Hellman, M Vanhove, A Lejeune, F J Stevens, and L M Hendershot. The in vivo association of bip with newly synthesized proteins is dependent on the rate and stability of folding and not simply on the presence of sequences that can bind to bip. *J Cell Biol*, 144(1):21–30, Jan 1999.

[201] Mats Hellström, Li-Kun Phng, and Holger Gerhardt. Vegf and notch signaling: the yin and yang of angiogenic sprouting. *Cell Adh Migr*, 1(3):133–6, 2007.

[202] Mats Hellström, Li-Kun Phng, Jennifer J Hofmann, Elisabet Wallgard, Leigh Coultas, Per Lindblom, Jackelyn Alva, Ann-Katrin Nilsson, Linda Karlsson, Nicholas Gaiano, Keejung Yoon, Janet Rossant, M Luisa Iruela-Arispe, Mattias Kalén, Holger Gerhardt, and Christer Betsholtz. Dll4 signalling through notch1 regulates formation of tip cells during angiogenesis. *Nature*, 445(7129):776–80, Feb 2007.

[203] Mohamed Helmy, Jin Gohda, Jun-Ichiro Inoue, Masaru Tomita, Masa Tsuchiya, and Kumar Selvarajoo. Predicting novel features of toll-like receptor 3 signaling in macrophages. *PLoS One*, 4(3):e4661, 2009.

[204] C. Hermann, B. Assmus, C. Urbich, A. M. Zeiher, and S. Dimmeler. Insulin-mediated stimulation of protein kinase Akt: A potent survival signaling cascade for endothelial cells. *Arterioscler Thromb Vasc Biol*, 20(2):402–409, Feb 2000.

[205] A Hershko, A Ciechanover, and A Varshavsky. Basic medical research award. the ubiquitin system. *Nat Med*, 6(10):1073–81, Oct 2000.

[206] C. Hetz, P. Bernasconi, J. Fisher, A.H. Lee, M.C. Bassik, B. Antonsson, G.S. Brandt, N.N. Iwakoshi, A. Schinzel, L.H. Glimcher, et al. Proapoptotic BAX and BAK modulate the unfolded protein response by a direct interaction with IRE1α, 2006.

[207] Claudio Hetz, Paula Bernasconi, Jill Fisher, Ann-Hwee Lee, Michael C Bassik, Bruno Antonsson, Gabriel S Brandt, Neal N Iwakoshi, Anna Schinzel, Laurie H Glimcher, and Stanley J Korsmeyer. Proapoptotic bax and bak modulate the unfolded protein response by a direct interaction with ire1alpha. *Science*, 312(5773):572–6, Apr 2006.

[208] Claudio Hetz and Laurie H Glimcher. Fine-tuning of the unfolded protein response: Assembling the ire1alpha interactome. *Mol Cell*, 35:551–61, Sep 2009.

[209] Steven M Hill, David E Blask, Shulin Xiang, Lin Yuan, Lulu Mao, Robert T Dauchy, Erin M Dauchy, Tripp Frasch, and Tamika Duplesis. Melatonin and associated signaling pathways that control normal breast epithelium and breast cancer. *J Mammary Gland Biol Neoplasia*, 16(3):235–45, Sep 2011.

[210] T Hiroi, H Wei, C Hough, P Leeds, and D-M Chuang. Protracted lithium treatment protects against the er stress elicited by thapsigargin in rat pc12 cells: roles of intracellular calcium, grp78 and bcl-2. *Pharmacogenomics J*, 5(2):102–11, 2005.

[211] William S Hlavacek, James R Faeder, Michael L Blinov, Alan S Perelson, and Byron Goldstein. The complexity of complexes in signal transduction. *Biotechnol Bioeng*, 84(7):783–94, Dec 2003.

[212] M Hockel, K Schlenger, B Aral, M Mitze, U Schaffer, and P Vaupel. Association between tumor hypoxia and malignant progression in advanced cancer of the uterine cervix. *Cancer Res*, 56(19):4509–15, Oct 1996.

[213] Thomas Holbro and Nancy E Hynes. Erbb receptors: directing key signaling networks throughout life. *Annu Rev Pharmacol Toxicol*, 44:195–217, 2004.

[214] Marina K Holz, Bryan A Ballif, Steven P Gygi, and John Blenis. mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. *Cell*, 123(4):569–580, Nov 2005.

[215] Min Hong, Shengzhan Luo, Peter Baumeister, Jen-Ming Huang, Raveen K Gogia, Mingqing Li, and Amy S Lee. Underglycosylation of atf6 as a novel sensing mechanism for activation of the unfolded protein response. *J Biol Chem*, 279(12):11354–63, Mar 2004.

[216] J J M Hoozemans, R Veerhuis, E S Van Haastert, J M Rozemuller, F Baas, P Eikelenboom, and W Scheper. The unfolded protein response is activated in alzheimer's disease. *Acta Neuropathol*, 110(2):165–72, Aug 2005.

[217] N Hosokawa, I Wada, K Hasegawa, T Yorihuzi, L O Tremblay, A Herscovics, and K Nagata. A novel er alpha-mannosidase-like protein accelerates er-associated degradation. *EMBO Rep*, 2(5):415–22, May 2001.

[218] Maria Høyer-Hansen, Lone Bastholm, Piotr Szyniarowski, Michelangelo Campanella, György Szabadkai, Thomas Farkas, Katiuscia Bianchi, Nicole Fehrenbacher, Folmer Elling, Rosario Rizzuto, Ida Stenfeldt Mathiasen, and Marja Jäättelä. Control of macroautophagy by calcium, calmodulin-dependent kinase kinase-beta, and bcl-2. *Mol Cell*, 25(2):193–205, Jan 2007.

[219] P. Hu, Z. Han, A.D. Couvillon, R.J. Kaufman, and J.H. Exton. Autocrine tumor necrosis factor alpha links endoplasmic reticulum stress to the membrane death receptor pathway through IRE1α-mediated NF-κB activation and down-regulation of TRAF2 expression. *Mol. Cell. Biol.*, 26(8):3071–3084, 2006.

[220] David T Hughes and Vanessa Sperandio. Inter-kingdom signalling: communication between bacteria and their hosts. *Nat Rev Microbiol*, 6(2):111–20, Feb 2008.

[221] Daniel R Hyduke and Bernhard Ø Palsson. Towards genome-scale signalling network reconstructions. *Nat Rev Genet*, 11(4):297–307, Apr 2010.

[222] Nancy E Hynes and Gwen MacDonald. Erbb receptors and signaling pathways in cancer. *Curr Opin Cell Biol*, 21(2):177–84, Apr 2009.

[223] Sergio Iadevaia, Yiling Lu, Fabiana C Morales, Gordon B Mills, and Prahlad T Ram. Identification of optimal drug combinations targeting cellular networks: integrating phospho-proteomics and computational network analysis. *Cancer Res*, 70(17):6704–6714, Sep 2010.

[224] Y Imai, M Soda, and R Takahashi. Parkin suppresses unfolded protein

stress-induced cell death through its e3 ubiquitin-protein ligase activity. *J Biol Chem*, 275(46):35661–4, Nov 2000.

[225] Ken Inoki, Yong Li, Tianquan Zhu, Jun Wu, and Kun-Liang Guan. TSC2 is phosphorylated and inhibited by Akt and suppresses mTOR signalling. *Nat Cell Biol*, 4(9):648–657, Sep 2002.

[226] Rafael A Irizarry, Bridget Hobbs, Francois Collin, Yasmin D Beazer-Barclay, Kristen J Antonellis, Uwe Scherf, and Terence P Speed. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4(2):249–64, Apr 2003.

[227] Hisamitsu Ishihara, Satoshi Takeda, Akira Tamura, Rui Takahashi, Suguru Yamaguchi, Daisuke Takei, Takahiro Yamada, Hiroshi Inoue, Hiroyuki Soga, Hideki Katagiri, Yukio Tanizawa, and Yoshitomo Oka. Disruption of the wfs1 gene in mice causes progressive beta-cell loss and impaired stimulus-secretion coupling in insulin secretion. *Hum Mol Genet*, 13(11):1159–70, Jun 2004.

[228] Richard J Jackson, Christopher U T Hellen, and Tatyana V Pestova. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol*, 11(2):113–127, Feb 2010.

[229] Lars Jakobsson, Claudio A Franco, Katie Bentley, Russell T Collins, Bas Ponsioen, Irene M Aspalter, Ian Rosewell, Marta Busse, Gavin Thurston, Alexander Medvinsky, Stefan Schulte-Merker, and Holger Gerhardt. Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. *Nat Cell Biol*, 12(10):943–53, Oct 2010.

[230] C Jamora, G Dennert, and A S Lee. Inhibition of tumor progression by suppression of stress protein grp78/bip induction in fibrosarcoma b/c10me. *Proc Natl Acad Sci U S A*, 93(15):7690–4, Jul 1996.

[231] Delphine Javelaud and Alain Mauviel. Crosstalk mechanisms between the mitogen-activated protein kinase pathways and smad signaling downstream of tgf-beta: implications for carcinogenesis. *Oncogene*, 24(37):5742–50, Aug 2005.

[232] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman. Global cancer statistics. *CA Cancer J Clin*, 61(2):69–90, 2011.

[233] Lars J Jensen, Michael Kuhn, Manuel Stark, Samuel Chaffron, Chris

Creevey, Jean Muller, Tobias Doerks, Philippe Julien, Alexander Roth, Milan Simonovic, Peer Bork, and Christian von Mering. String 8–a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res*, 37(Database issue):D412–6, Jan 2009.

[234] H Jeong, S P Mason, A L Barabási, and Z N Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41–2, May 2001.

[235] H. Jiang and P. K. Vogt. Constitutively active Rheb induces oncogenic transformation. *Oncogene*, 27(43):5729–5740, Sep 2008.

[236] A. Jimbo, E. Fujita, Y. Kouroku, J. Ohnishi, N. Inohara, K. Kuida, K. Sakamaki, S. Yonehara, and T. Momoi. ER stress induces caspase-8 activation, stimulating cytochrome c release and caspase-9 activation. *Experimental cell research*, 283(2):156–166, 2003.

[237] Ingunn W Jolma, Ole Didrik Laerum, Cathrine Lillo, and Peter Ruoff. Circadian oscillators in eukaryotes. *Wiley Interdiscip Rev Syst Biol Med*, 2(5):533–49, 2010.

[238] Chihiro Kakiuchi, Mizuho Ishiwata, Shinichiro Nanko, Hiroshi Kunugi, Yoshio Minabe, Kazuhiko Nakamura, Norio Mori, Kumiko Fujii, Tadashi Umekage, Mamoru Tochigi, Kazuhisa Kohda, Tsukasa Sasaki, Kazuo Yamada, Takeo Yoshikawa, and Tadafumi Kato. Functional polymorphisms of hspa5: possible association with bipolar disorder. *Biochem Biophys Res Commun*, 336(4):1136–43, Nov 2005.

[239] Chihiro Kakiuchi, Kazuya Iwamoto, Mizuho Ishiwata, Miki Bundo, Takaoki Kasahara, Ichiro Kusumi, Takahiro Tsujita, Yuji Okazaki, Shinichiro Nanko, Hiroshi Kunugi, Tsukasa Sasaki, and Tadafumi Kato. Impaired feedback regulation of xbp1 as a genetic risk factor for bipolar disorder. *Nat Genet*, 35(2):171–5, Oct 2003.

[240] Takahiro Kamimoto, Shisako Shoji, Tunda Hidvegi, Noboru Mizushima, Kyohei Umebayashi, David H Perlmutter, and Tamotsu Yoshimori. Intracellular inclusions containing mutant alpha1-antitrypsin z are propagated in the absence of autophagic activity. *J Biol Chem*, 281(7):4467–76, Feb 2006.

[241] M Karakurum, R Shreeniwas, J Chen, D Pinsky, S D Yan, M Anderson, K Sunouchi, J Major, T Hamilton, and K Kuwabara. Hypoxic induction of interleukin-8 gene expression in human endothelial cells. *J Clin Invest*, 93(4):1564–70, Apr 1994.

[242] T Katayama, K Imaizumi, N Sato, K Miyoshi, T Kudo, J Hitomi, T Morihara, T Yoneda, F Gomi, Y Mori, Y Nakano, J Takeda, T Tsuda, Y Itoyama, O Murayama, A Takashima, P St George-Hyslop, M Takeda, and M Tohyama. Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. *Nat Cell Biol*, 1(8):479–85, Dec 1999.

[243] Taiichi Katayama, Kazunori Imaizumi, Takayuki Manabe, Junichi Hitomi, Takashi Kudo, and Masaya Tohyama. Induction of neuronal death by er stress in alzheimer's disease. *J Chem Neuroanat*, 28(1-2):67–78, Sep 2004.

[244] Masuko Katoh and Masaru Katoh. Cross-talk of wnt and fgf signaling pathways at gsk3beta to regulate beta-catenin and snail signaling cascades. *Cancer Biol Ther*, 5(9):1059–64, Sep 2006.

[245] Randal J Kaufman, Donalyn Scheuner, Martin Schröder, Xiaohua Shen, Kyungho Lee, Chuan Yin Liu, and Stacey M Arnold. The unfolded protein response in nutrient sensing and differentiation. *Nat Rev Mol Cell Biol*, 3(6):411–21, Jun 2002.

[246] Sem Kebache, Eric Cardin, Duc Thang Nguyên, Eric Chevet, and Louise Larose. Nck-1 antagonizes the endoplasmic reticulum stress-induced inhibition of translation. *J Biol Chem*, 279(10):9662–71, Mar 2004.

[247] J F Kerr, A H Wyllie, and A R Currie. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer*, 26(4):239–57, Aug 1972.

[248] David Kessel. Protection of bcl-2 by salubrinal. *Biochem Biophys Res Commun*, 346(4):1320–3, Aug 2006.

[249] A Khanna and R D Campbell. The gene g13 in the class iii region of the human mhc encodes a potential dna-binding protein. *Biochem J*, 319 (Pt 1):81–9, Oct 1996.

[250] A H Kim, G Khursigara, X Sun, T F Franke, and M V Chao. Akt phosphorylates and negatively regulates apoptosis signal-regulating kinase 1. *Mol Cell Biol*, 21(3):893–901, Feb 2001.

[251] Hyung-Do Kim, Aaron S Meyer, Joel P Wagner, Shannon K Alford, Alan Wells, Frank B Gertler, and Douglas A Lauffenburger. Signaling network state predicts twist-mediated effects on breast cell migration across diverse growth factor contexts. *Mol Cell Proteomics*, Aug 2011.

[252] Hyunsoo Kim, Haesun Park, and Barry L Drake. Extracting unrecognized gene relationships from the biomedical literature via matrix factorizations. *BMC Bioinformatics*, 8 Suppl 9:S6, 2007.

[253] Philip M Kim, Andrea Sboner, Yu Xia, and Mark Gerstein. The role of disorder in interaction networks: a structural analysis. *Mol Syst Biol*, 4:179, 2008.

[254] Yukio Kimata, Yuki Ishiwata-Kimata, Tatsuhiko Ito, Aiko Hirata, Tomohide Suzuki, Daisuke Oikawa, Masato Takeuchi, and Kenji Kohno. Two regulatory steps of er-stress sensor ire1 involving its cluster formation and interaction with unfolded proteins. *J Cell Biol*, 179(1):75–86, Oct 2007.

[255] Yukio Kimata, Yuki I Kimata, Yusuke Shimizu, Hiroshi Abe, Ileana C Farcasanu, Masato Takeuchi, Mark D Rose, and Kenji Kohno. Genetic evidence for a role of bip/kar2 that regulates ire1 in response to accumulation of unfolded proteins. *Mol Biol Cell*, 14(6):2559–69, Jun 2003.

[256] H. Kitano. A robustness based approach to systems-oriented drug design. *Nat. Rev. Drug Discov.*, 6:202 – 210, 2007.

[257] Hiroaki Kitano. Computational systems biology. *Nature*, 420(6912):206–10, Nov 2002.

[258] Hiroaki Kitano. Systems biology: a brief overview. *Science*, 295(5560):1662–4, Mar 2002.

[259] Hiroaki Kitano. Cancer as a robust system: implications for anticancer therapy. *Nat Rev Cancer*, 4(3):227–35, Mar 2004.

[260] K Kitchin and M C Flickinger. Alteration of hybridoma viability and antibody secretion in transfectomas with inducible overexpression of protein disulfide isomerase. *Biotechnol Prog*, 11(5):565–74, 1995.

[261] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International joint Conference on artificial intelligence*, volume 14, pages 1137–1145. Citeseer, 1995.

[262] K Kohno, K Normington, J Sambrook, M J Gething, and K Mori. The promoter region of the yeast kar2 (bip) gene contains a regulatory domain that responds to the presence of unfolded proteins in the endoplasmic reticulum. *Mol Cell Biol*, 13(2):877–90, Feb 1993.

[263] K Kokame, H Kato, and T Miyata. Identification of erse-ii, a new cis-acting element responsible for the atf6-dependent mammalian unfolded protein response. *J Biol Chem*, 276(12):9199–205, Mar 2001.

[264] A C Koong, E Y Chen, and A J Giaccia. Hypoxia causes the activation of nuclear factor kappa b through the phosphorylation of i kappa b alpha on tyrosine residues. *Cancer Res*, 54(6):1425–30, Mar 1994.

[265] Marianne Koritzinsky, Michaël G Magagnin, Twan van den Beucken, Renaud Seigneuric, Kim Savelkouls, Josée Dostie, Stéphane Pyronnet, Randal J Kaufman, Sherry A Weppler, Jan Willem Voncken, Philippe Lambin, Constantinos Koumenis, Nahum Sonenberg, and Bradly G Wouters. Gene expression during acute and prolonged hypoxia is regulated by distinct mechanisms of translational control. *EMBO J*, 25(5):1114–25, Mar 2006.

[266] Markus Koschorreck, Holger Conzelmann, Sybille Ebert, Michael Ederer, and Ernst Dieter Gilles. Reduced modeling of signal transduction - a modular approach. *BMC Bioinformatics*, 8:336, 2007.

[267] Shilpa Kothari, Jeannick Cizeau, Eileen McMillan-Ward, Sara J Israels, Michelle Bailes, Karen Ens, Lorrie A Kirshenbaum, and Spencer B Gibson. Bnip3 plays a role in hypoxic cell death in human epithelial cells that is inhibited by growth factors egf and igf. *Oncogene*, 22(30):4734–44, Jul 2003.

[268] Constantinos Koumenis, Christine Naczki, Marianne Koritzinsky, Sally Rastani, Alan Diehl, Nahum Sonenberg, Antonis Koromilas, and Bradly G Wouters. Regulation of protein synthesis by hypoxia via activation of the endoplasmic reticulum kinase perk and phosphorylation of the translation initiation factor eif2alpha. *Mol Cell Biol*, 22(21):7405–16, Nov 2002.

[269] Constantinos Koumenis and Bradly G Wouters. "translating" tumor hypoxia: unfolded protein response (upr)-dependent and upr-independent pathways. *Mol Cancer Res*, 4(7):423–36, Jul 2006.

[270] Y Kouroku, E Fujita, I Tanida, T Ueno, A Isoai, H Kumagai, S Ogawa, R J Kaufman, E Kominami, and T Momoi. Er stress (perk/eif2alpha phosphorylation) mediates the polyglutamine-induced lc3 conversion, an essential step for autophagy formation. *Cell Death Differ*, 14(2):230–9, Feb 2007.

[271] Marcin Kowanetz and Napoleone Ferrara. Vascular endothelial growth factor signaling pathways: therapeutic perspective. *Clin Cancer Res*, 12(17):5018–22, Sep 2006.

[272] Pamela K Kreeger and Douglas A Lauffenburger. Cancer systems biology: a network modeling perspective. *Carcinogenesis*, 31(1):2–8, Jan 2010.

[273] A Kremling and J Saez-Rodriguez. Systems biology–an engineering perspective. *J Biotechnol*, 129(2):329–51, Apr 2007.

[274] Andreas Kremling, Sophia Fischer, Kapil Gadkar, Francis J Doyle, Thomas Sauter, Eric Bullinger, Frank Allgöwer, and Ernst D Gilles. A benchmark for methods in reverse engineering and model discrimination: problem formulation and solutions. *Genome Res*, 14(9):1773–85, Sep 2004.

[275] M Kretzschmar, J Doody, I Timokhina, and J Massagué. A mechanism of repression of tgfbeta/ smad signaling by oncogenic ras. *Genes Dev*, 13(7):804–16, Apr 1999.

[276] Guido Kroemer and Marja Jäättelä. Lysosomes and autophagy in cell death control. *Nat Rev Cancer*, 5(11):886–97, Nov 2005.

[277] Nevan J Krogan, Gerard Cagney, Haiyuan Yu, Gouqing Zhong, Xinghua Guo, Alexandr Ignatchenko, Joyce Li, Shuye Pu, Nira Datta, Aaron P Tikuisis, Thanuja Punna, José M Peregrín-Alvarez, Michael Shales, Xin Zhang, Michael Davey, Mark D Robinson, Alberto Paccanaro, James E Bray, Anthony Sheung, Bryan Beattie, Dawn P Richards, Veronica Canadien, Atanas Lalev, Frank Mena, Peter Wong, Andrei Starostine, Myra M Canete, James Vlasblom, Samuel Wu, Chris Orsi, Sean R Collins, Shamanta Chandran, Robin Haw, Jennifer J Rilstone, Kiran Gandi, Natalie J Thompson, Gabe Musso, Peter St Onge, Shaun Ghanny, Mandy H Y Lam, Gareth Butland, Amin M Altaf-Ul, Shigehiko Kanaya, Ali Shilatifard, Erin O'Shea, Jonathan S Weissman, C James Ingles, Timothy R Hughes, John Parkinson, Mark Gerstein, Shoshana J Wodak, Andrew Emili, and Jack F Greenblatt. Global landscape of protein complexes in the yeast saccharomyces cerevisiae. *Nature*, 440(7084):637–43, Mar 2006.

[278] Sebastian C Y Ku, Daphne T W Ng, Miranda G S Yap, and Sheng-Hao Chao. Effects of overexpression of x-box binding protein 1 on recombinant protein production in chinese hamster ovary and ns0 myeloma cells. *Biotechnol Bioeng*, 99(1):155–64, Jan 2008.

[279] T Kudo, S Kanemoto, H Hara, N Morimoto, T Morihara, R Kimura, T Tabira, K Imaizumi, and M Takeda. A molecular chaperone inducer protects neurons from er stress. *Cell Death Differ*, 15(2):364–75, Feb 2008.

[280] L. Kuepfer, M. Peter, U Sauer, and J. Stelling. Ensemble modeling for analysis of cell signaling dynamics. *Nat Biotech*, 25(9):1001–1006, 2007.

[281] R Kumar, S Azam, J M Sullivan, C Owen, D R Cavener, P Zhang, D Ron, H P Harding, J J Chen, A Han, B C White, G S Krause, and D J DeGracia. Brain ischemia and reperfusion activates the eukaryotic initiation factor 2alpha kinase, perk. *J Neurochem*, 77(5):1418–21, Jun 2001.

[282] Frank M LaFerla. Calcium dyshomeostasis and intracellular signalling in alzheimer's disease. *Nat Rev Neurosci*, 3(11):862–72, Nov 2002.

[283] Esther Latres, Ami R Amini, Ashley A Amini, Jennifer Griffiths, Francis J Martin, Yi Wei, Hsin Chieh Lin, George D Yancopoulos, and David J Glass. Insulin-like growth factor-1 (IGF-1) inversely regulates atrophy-induced genes via the phosphatidylinositol 3-kinase/Akt/mammalian target of rapamycin (PI3K/Akt/mTOR) pathway. *J Biol Chem*, 280(4):2737–2744, Jan 2005.

[284] Alexis Laurent, Carole Nicco, Christiane Chéreau, Claire Goulvestre, Jérôme Alexandre, Arnaud Alves, Eva Lévy, Francois Goldwasser, Yves Panis, Olivier Soubrane, Bernard Weill, and Frédéric Batteux. Controlling tumor growth by modulating endogenous production of reactive oxygen species. *Cancer Res*, 65(3):948–56, Feb 2005.

[285] Inna N Lavrik. Systems biology of apoptosis signaling networks. *Curr Opin Biotechnol*, 21(4):551–5, Aug 2010.

[286] Inna N Lavrik, Roland Eils, Nicolai Fricker, Carina Pforr, and Peter H Krammer. Understanding apoptosis by systems biology approaches. *Mol Biosyst*, 5(10):1105–11, Oct 2009.

[287] Matthew J Lazzara and Douglas A Lauffenburger. Quantitative modeling perspectives on the erbb system of cell regulatory processes. *Exp Cell Res*, 315(4):717–25, Feb 2009.

[288] Ann-Hwee Lee, Neal N Iwakoshi, Kenneth C Anderson, and Laurie H Glimcher. Proteasome inhibitors disrupt the unfolded protein response in myeloma cells. *Proc Natl Acad Sci U S A*, 100(17):9946–51, Aug 2003.

[289] Ann-Hwee Lee, Erez F Scapa, David E Cohen, and Laurie H Glimcher. Regulation of hepatic lipogenesis by the transcription factor xbp1. *Science*, 320(5882):1492–6, Jun 2008.

[290] Jonathan M Lee, Shoukat Dedhar, Raghu Kalluri, and Erik W Thompson. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *J Cell Biol*, 172(7):973–81, Mar 2006.

[291] Kyungho Lee, Witoon Tirasophon, Xiaohua Shen, Marek Michalak, Ron Prywes, Tetsuya Okada, Hiderou Yoshida, Kazutoshi Mori, and Randal J Kaufman. Ire1-mediated unconventional mrna splicing and s2p-mediated atf6 cleavage merge to regulate xbp1 in signaling the unfolded protein response. *Genes Dev*, 16(4):452–66, Feb 2002.

[292] T G Lee, J Tomita, A G Hovanessian, and M G Katze. Purification and partial characterization of a cellular inhibitor of the interferon-induced protein kinase of mr 68,000 from influenza virus-infected cells. *Proc Natl Acad Sci U S A*, 87(16):6208–12, Aug 1990.

[293] Tong Ihn Lee, Nicola J Rinaldi, François Robert, Duncan T Odom, Ziv Bar-Joseph, Georg K Gerber, Nancy M Hannett, Christopher T Harbison, Craig M Thompson, Itamar Simon, Julia Zeitlinger, Ezra G Jennings, Heather L Murray, D Benjamin Gordon, Bing Ren, John J Wyrick, Jean-Bosco Tagne, Thomas L Volkert, Ernest Fraenkel, David K Gifford, and Richard A Young. Transcriptional regulatory networks in saccharomyces cerevisiae. *Science*, 298(5594):799–804, Oct 2002.

[294] Kui Lei and Roger J Davis. Jnk phosphorylation of bim-related members of the bcl2 family induces bax-dependent apoptosis. *Proc Natl Acad Sci U S A*, 100(5):2432–7, Mar 2003.

[295] Beth Levine and Daniel J Klionsky. Development by self-digestion: molecular mechanisms and biological functions of autophagy. *Dev Cell*, 6(4):463–77, Apr 2004.

[296] Beth Levine and Guido Kroemer. Autophagy in the pathogenesis of disease. *Cell*, 132(1):27–42, Jan 2008.

[297] Hu Li, Choong Yong Ung, Xiao Hua Ma, Xiang Hui Liu, Bao Wen Li, Boon Chuan Low, and Yu Zong Chen. Pathway sensitivity analysis for detecting pro-proliferation activities of oncogenes and tumor suppressors of epidermal growth factor receptor-extracellular signal-regulated protein kinase pathway at altered protein levels. *Cancer*, 115(18):4246–4263, Sep 2009.

[298] J. Li, K. DeFea, and R. A. Roth. Modulation of insulin receptor substrate-1

tyrosine phosphorylation by an Akt/phosphatidylinositol 3-kinase pathway. *J Biol Chem*, 274(14):9351–9356, Apr 1999.

[299] Song Li, Jianmin Lao, Benjamin P C Chen, Yi-shuan Li, Yihua Zhao, Julia Chu, Kuang-Den Chen, Tsui-Chun Tsou, Konan Peck, and Shu Chien. Genomic analysis of smooth muscle cells in 3-dimensional collagen matrix. *FASEB J*, 17(1):97–9, Jan 2003.

[300] Debbie Liao, Courtney Corle, Tiffany N Seagroves, and Randall S Johnson. Hypoxia-inducible factor-1alpha is a key regulator of metastasis in a transgenic model of cancer initiation and progression. *Cancer Res*, 67(2):563–72, Jan 2007.

[301] J.H. Lin, H. Li, D. Yasumura, H.R. Cohen, C. Zhang, B. Panning, K.M. Shokat, M.M. LaVail, and P. Walter. IRE1 signaling affects cell fate during the unfolded protein response. *Science*, 318(5852):944, 2007.

[302] J.H. Lin, H. Li, Y. Zhang, D. Ron, and P. Walter. Divergent effects of PERK and IRE1 signaling on cell viability. *PLoS One*, 4(1):4170, 2009.

[303] Li-Hsieh Lin, Hsiao-Ching Lee, Wen-Hsiung Li, and Bor-Sen Chen. A systematic approach to detecting transcription factors in response to environmental stresses. *BMC Bioinformatics*, 8:473, 2007.

[304] D Lindholm, H Wootz, and L Korhonen. Er stress and neurodegenerative diseases. *Cell Death Differ*, 13(3):385–92, Mar 2006.

[305] Rune Linding, Lars Juhl Jensen, Gerard J Ostheimer, Marcel A T M van Vugt, Claus Jørgensen, Ioana M Miron, Francesca Diella, Karen Colwill, Lorne Taylor, Kelly Elder, Pavel Metalnikov, Vivian Nguyen, Adrian Pasculescu, Jing Jin, Jin Gyoon Park, Leona D Samson, James R Woodgett, Robert B Russell, Peer Bork, Michael B Yaffe, and Tony Pawson. Systematic discovery of in vivo phosphorylation networks. *Cell*, 129(7):1415–26, Jun 2007.

[306] M. N. Lioubin, P. A. Algate, S. Tsai, K. Carlberg, A. Aebersold, and L. R. Rohrschneider. p150Ship, a signal transduction molecule with inositol polyphosphate-5-phosphatase activity. *Genes Dev*, 10(9):1084–1095, May 1996.

[307] Fernanda Lisbona, Diego Rojas-Rivera, Peter Thielen, Sebastian Zamorano, Derrick Todd, Fabio Martinon, Alvaro Glavic, Christina Kress,

Jonathan H Lin, Peter Walter, John C Reed, Laurie H Glimcher, and Claudio Hetz. Bax inhibitor-1 is a negative regulator of the er stress sensor ire1alpha. *Mol Cell*, 33(6):679–91, Mar 2009.

[308] Chuan Yin Liu, Hetty N Wong, Joseph A Schauerte, and Randal J Kaufman. The protein kinase/endoribonuclease ire1alpha that signals the unfolded protein response has a luminal n-terminal ligand-independent dimerization domain. *J Biol Chem*, 277(21):18346–56, May 2002.

[309] Chuan Yin Liu, Zhaohui Xu, and Randal J Kaufman. Structure and intermolecular interactions of the luminal dimerization domain of human ire1alpha. *J Biol Chem*, 278(20):17680–7, May 2003.

[310] Gang Liu, Amina A Qutub, Prakash Vempati, Feilim Mac Gabhann, and Aleksander S Popel. Module-based multiscale simulation of angiogenesis in skeletal muscle. *Theor Biol Med Model*, 8:6, 2011.

[311] Hong Liu, Derek C Radisky, Celeste M Nelson, Hui Zhang, Jimmie E Fata, Richard A Roth, and Mina J Bissell. Mechanism of Akt1 inhibition of breast cancer cell invasion reveals a protumorigenic role for TSC2. *Proc Natl Acad Sci U S A*, 103(11):4134–4139, Mar 2006.

[312] L Liu, X Ning, S Han, H Zhang, L Sun, Y Shi, S Sun, C Guo, F Yin, T Qiao, K Wu, and D Fan. [hypoxia induced hif-1 accumulation and vegf expression in gastric epithelial mucosa cell: involvement of erk1/2 and pi3k/akt]. *Mol Biol (Mosk)*, 42(3):459–69, 2008.

[313] Y Liu, P Choudhury, C M Cabral, and R N Sifers. Oligosaccharide modification in the early secretory pathway directs the selection of a misfolded glycoprotein for degradation by the proteasome. *J Biol Chem*, 274(9):5861–7, Feb 1999.

[314] Yingting Liu, Jeremy Purvis, Andrew Shih, Joshua Weinstein, Neeraj Agrawal, and Ravi Radhakrishnan. A multiscale computational approach to dissect early events in the erb family receptor mediated activation, differential signaling, and relevance to oncogenic transformations. *Ann Biomed Eng*, 35(6):1012–25, Jun 2007.

[315] Agnieszka Loboda, Anna Stachurska, Urszula Florczyk, Dominika Rudnicka, Agnieszka Jazwa, Joanna Wegrzyn, Magdalena Kozakowska, Krystyna Stalinska, Lorenz Poellinger, Anna-Liisa Levonen, Seppo Yla-Herttuala, Alicja Jozkowicz, and Jozef Dulak. Hif-1 induction attenuates

nrf2-dependent il-8 expression in human endothelial cells. *Antioxid Redox Signal*, 11(7):1501–17, Jul 2009.

[316] J. R. Lorsch and D. Herschlag. Kinetic dissection of fundamental processes of eukaryotic translation initiation in vitro. *EMBO J*, 18(23):6705–6717, Dec 1999.

[317] Robyn M B Loureiro, Arindel S R Maharaj, David Dankort, William J Muller, and Patricia A D'Amore. Erbb2 overexpression in mammary cells upregulates vegf through the core promoter. *Biochem Biophys Res Commun*, 326(2):455–65, Jan 2005.

[318] Phoebe D Lu, Heather P Harding, and David Ron. Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J Cell Biol*, 167(1):27–33, Oct 2004.

[319] Deyan Luan, Fania Szlam, Kenichi A Tanaka, Philip S Barie, and Jeffrey D Varner. Ensembles of uncertain mathematical models can identify network response to therapeutic interventions. *Mol Biosyst*, 6(11):2272–86, Nov 2010.

[320] Deyan Luan, Michael Zai, and Jeffrey D Varner. Computationally derived points of fragility of a human cascade are consistent with current therapeutic strategies. *PLoS Comput Biol*, 3(7):e142, Jul 2007.

[321] Dianhong Luo, Yun He, Haifeng Zhang, Luyang Yu, Hong Chen, Zhe Xu, Shibo Tang, Fumihiko Urano, and Wang Min. Aip1 is critical in transducing ire1-mediated endoplasmic reticulum stress response. *J Biol Chem*, 283(18):11905–12, May 2008.

[322] Xiaoju Max Ma and John Blenis. Molecular mechanisms of mTOR-mediated translational control. *Nat Rev Mol Cell Biol*, 10(5):307–318, May 2009.

[323] Xiaoju Max Ma, Sang-Oh Yoon, Celeste J Richardson, Kristina Julich, and John Blenis. SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. *Cell*, 133(2):303–313, Apr 2008.

[324] Yanjun Ma and Linda M Hendershot. Delineation of a negative feedback regulatory loop that controls protein translation during endoplasmic reticulum stress. *J Biol Chem*, 278(37):34864–73, Sep 2003.

[325] Pekka Määttänen, Kalle Gehring, John J M Bergeron, and David Y Thomas. Protein quality control in the er: the recognition of misfolded proteins. *Semin Cell Dev Biol*, 21(5):500–11, Jul 2010.

[326] Feilim Mac Gabhann, James W Ji, and Aleksander S Popel. Computational model of vascular endothelial growth factor spatial distribution in muscle and pro-angiogenic cell therapy. *PLoS Comput Biol*, 2(9):e127, Sep 2006.

[327] Colin Macilwain. Systems biology: evolving into the mainstream. *Cell*, 144(6):839–41, Mar 2011.

[328] P. Macklin, J. Kim, G. Tomaiuolo, M.E. Edgerton, and V. Cristini. Agent-based modeling of ductal carcinoma in situ: application to patient-specific breast cancer modeling. *Computational Biology*, pages 77–111, 2010.

[329] Shyuichi Maeda, Kazuyuki Nakajima, Yoko Tohyama, and Shinichi Kohsaka. Characteristic response of astrocytes to plasminogen/plasmin to upregulate transforming growth factor beta 3 (tgfbeta3) production/secretion through proteinase-activated receptor-1 (par-1) and the downstream phosphatidylinositol 3-kinase (pi3k)-akt/pkb signaling cascade. *Brain Res*, 1305:1–13, Dec 2009.

[330] Jyoti D Malhotra and Randal J Kaufman. The endoplasmic reticulum and the unfolded protein response. *Semin Cell Dev Biol*, 18(6):716–31, Dec 2007.

[331] Brendan D Manning, Andrew R Tee, M. Nicole Logsdon, John Blenis, and Lewis C Cantley. Identification of the tuberous sclerosis complex-2 tumor suppressor gene product tuberin as a target of the phosphoinositide 3-kinase/akt pathway. *Mol Cell*, 10(1):151–162, Jul 2002.

[332] Monica G Marcu, Melissa Doyle, Anne Bertolotti, David Ron, Linda Hendershot, and Len Neckers. Heat shock protein 90 modulates the unfolded protein response by stabilizing ire1alpha. *Mol Cell Biol*, 22(24):8506–13, Dec 2002.

[333] Assen Marintchev and Gerhard Wagner. Translation initiation: structures, mechanisms and evolution. *Q Rev Biophys*, 37(3-4):197–284, 2004.

[334] Suchita Markan, Harbir Singh Kohli, Kusum Joshi, Ranjana Walker Minz, Kamal Sud, Monika Ahuja, Shashi Anand, and Madhu Khullar. Up regulation of the grp-78 and gadd-153 and down regulation of bcl-2 proteins in primary glomerular diseases: a possible involvement of the er stress pathway in glomerulonephritis. *Mol Cell Biochem*, 324(1-2):131–8, Apr 2009.

[335] Mina D Marmor, Kochupurakkal Bose Skaria, and Yosef Yarden. Signal transduction and oncogenesis by erbb/her receptors. *Int J Radiat Oncol Biol Phys*, 58(3):903–13, Mar 2004.

[336] Daniel Martin, Rebeca Galisteo, and J Silvio Gutkind. Cxcl8/il8 stimulates vascular endothelial growth factor (vegf) expression and the autocrine activation of vegfr2 in endothelial cells by activating nfkappab through the cbm (carma3/bcl10/malt1) complex. *J Biol Chem*, 284(10):6038–42, Mar 2009.

[337] J Massagué. How cells read tgf-beta signals. *Nat Rev Mol Cell Biol*, 1(3):169–78, Dec 2000.

[338] Joan Massague. Integration of smad and mapk pathways: a link and a linker revisited. *Genes Dev*, 17(24):2993–7, Dec 2003.

[339] Joan Massagué, Joan Seoane, and David Wotton. Smad transcription factors. *Genes Dev*, 19(23):2783–810, Dec 2005.

[340] K E Matlack, W Mothes, and T A Rapoport. Protein translocation: tunnel vision. *Cell*, 92(3):381–90, Feb 1998.

[341] M Matsumoto, M Minami, K Takeda, Y Sakao, and S Akira. Ectopic expression of chop (gadd153) induces apoptosis in m1 myeloblastic leukemia cells. *FEBS Lett*, 395:143–7, Oct 1996.

[342] M P Mattson. Apoptosis in neurodegenerative disorders. *Nat Rev Mol Cell Biol*, 1(2):120–9, Nov 2000.

[343] P J Maxwell, R Gallagher, A Seaton, C Wilson, P Scullin, J Pettigrew, I J Stratford, K J Williams, P G Johnston, and D J J Waugh. Hif-1 and nf-kappab-mediated upregulation of cxcr1 and cxcr2 expression promotes cell survival in hypoxic prostate cancer cells. *Oncogene*, 26(52):7333–45, Nov 2007.

[344] E V Maytin, M Ubeda, J C Lin, and J F Habener. Stress-inducible transcription factor chop/gadd153 induces apoptosis in mammalian cells via p38 kinase-dependent and -independent mechanisms. *Exp Cell Res*, 267:193–204, Jul 2001.

[345] K D McCullough, J L Martindale, L O Klotz, T Y Aw, and N J Holbrook. Gadd153 sensitizes cells to endoplasmic reticulum stress by down-

regulating bcl2 and perturbing the cellular redox state. *Mol Cell Biol*, 21(4):1249–59, Feb 2001.

[346] Damian Medici, Elizabeth D Hay, and Daniel A Goodenough. Cooperation between snail and lef-1 transcription factors is essential for tgf-beta1-induced epithelial-mesenchymal transition. *Mol Biol Cell*, 17(4):1871–9, Apr 2006.

[347] Damian Medici, Elizabeth D Hay, and Bjorn R Olsen. Snail and slug promote epithelial-mesenchymal transition through beta-catenin-t-cell factor-4-dependent expression of transforming growth factor-beta3. *Mol Biol Cell*, 19(11):4875–87, Nov 2008.

[348] Alfred J Meijer and Patrice Codogno. Signalling and autophagy regulation in health, aging and disease. *Mol Aspects Med*, 27(5-6):411–25, 2006.

[349] Giovanni Melillo. Targeting hypoxia cell signaling for cancer therapy. *Cancer Metastasis Rev*, 26(2):341–52, Jun 2007.

[350] Idan Menashe, Dennis Maeder, Montserrat Garcia-Closas, Jonine D Figueroa, Samsiddhi Bhattacharjee, Melissa Rotunno, Peter Kraft, David J Hunter, Stephen J Chanock, Philip S Rosenberg, and Nilanjan Chatterjee. Pathway analysis of breast cancer genome-wide association study highlights three pathways and one canonical signaling cascade. *Cancer Res*, 70(11):4453–9, Jun 2010.

[351] William C Merrick. Cap-dependent and cap-independent translation in eukaryotic systems. *Gene*, 332:1–11, 2004.

[352] Ollivier Milhavet, Jennifer L Martindale, Simonetta Camandola, Sic L Chan, Devin S Gary, Aiwu Cheng, Nikki J Holbrook, and Mark P Mattson. Involvement of gadd153 in the pathogenic action of presenilin-1 mutations. *J Neurochem*, 83(3):673–81, Nov 2002.

[353] R Milo, S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii, and U Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–7, Oct 2002.

[354] J Min, H Shukla, H Kozono, S K Bronson, S M Weissman, and D D Chaplin. A novel creb family gene telomeric of hla-dra in the hla complex. *Genomics*, 30(2):149–56, Nov 1995.

[355] Yusuke Mizukami, Yutaka Kohgo, and Daniel C Chung. Hypoxia inducible factor-1 independent pathways in tumor angiogenesis. *Clin Cancer Res*, 13(19):5670–4, Oct 2007.

[356] Chaya Mohan, Yeon-Gu Kim, Jane Koo, and Gyun Min Lee. Assessment of cell engineering strategies for improved therapeutic protein production in cho cells. *Biotechnol J*, 3(5):624–30, May 2008.

[357] Alfonso Mora, David Komander, Daan M F van Aalten, and Dario R Alessi. PDK1, the master regulator of AGC kinase signal transduction. *Semin Cell Dev Biol*, 15(2):161–170, Apr 2004.

[358] Katsutaro Morino, Kitt Falk Petersen, and Gerald I Shulman. Molecular mechanisms of insulin resistance in humans and their potential links with mitochondrial dysfunction. *Diabetes*, 55 Suppl 2:S9–S15, Dec 2006.

[359] K Morita, M Saitoh, K Tobiume, H Matsuura, S Enomoto, H Nishitoh, and H Ichijo. Negative feedback regulation of ask1 by protein phosphatase 5 (pp5) in response to oxidative stress. *EMBO J*, 20(21):6028–36, Nov 2001.

[360] M Muzio, B R Stockwell, H R Stennicke, G S Salvesen, and V M Dixit. An induced proximity model for caspase-8 activation. *J Biol Chem*, 273(5):2926–30, Jan 1998.

[361] M. G. Myers, R. Mendez, P. Shi, J. H. Pierce, R. Rhoads, and M. F. White. The COOH-terminal tyrosine phosphorylation sites on IRS-1 bind SHP-2 and negatively regulate insulin signaling. *J Biol Chem*, 273(41):26908–26914, Oct 1998.

[362] S. Nagl, M. Williams, and J. Williamson. Objective bayesian nets for systems modelling and prognosis in breast cancer. *Innovations in Bayesian Networks*, pages 131–167, 2008.

[363] Nirinjini Naidoo. Er and aging-protein folding and the er stress response. *Ageing Res Rev*, 8(3):150–9, Jul 2009.

[364] T Nakagawa, H Zhu, N Morishima, E Li, J Xu, B A Yankner, and J Yuan. Caspase-12 mediates endoplasmic-reticulum-specific apoptosis and cytotoxicity by amyloid-beta. *Nature*, 403(6765):98–103, Jan 2000.

[365] Jyoti Nautiyal, Pralay Majumder, Bhaumik B Patel, Francis Y Lee, and Adhip P N Majumdar. Src inhibitor dasatinib inhibits growth of breast

cancer cells by modulating egfr signaling. *Cancer Lett*, 283(2):143–51, Oct 2009.

[366] Steffan T Nawrocki, Jennifer S Carew, Kenneth Dunner, Jr, Lawrence H Boise, Paul J Chiao, Peng Huang, James L Abbruzzese, and David J McConkey. Bortezomib inhibits pkr-like endoplasmic reticulum (er) kinase and induces apoptosis via er stress in human pancreatic cancer cells. *Cancer Res*, 65(24):11510–9, Dec 2005.

[367] Ali Nawshad, Damian Medici, Chang-Chih Liu, and Elizabeth D Hay. Tgfbeta3 inhibits e-cadherin gene expression in palate medial-edge epithelial cells through a smad2-smad4-lef1 transcription complex. *J Cell Sci*, 120(Pt 9):1646–53, May 2007.

[368] S. Nayak, J. K. Siddiqui, and J. D. Varner. Modelling and analysis of an ensemble of eukaryotic translation initiation models. *IET Systems Biology*, 5(1):2–14, 2011.

[369] Satyaprakash Nayak, Saniya Salim, Deyan Luan, Michael Zai, and Jeffrey D Varner. A test of highly optimized tolerance reveals fragile cell-cycle mechanisms are molecular targets in clinical cancer trials. *PLoS One*, 3(4):e2016, 2008.

[370] A Nechushtan, C L Smith, Y T Hsu, and R J Youle. Conformation of the bax c-terminus regulates subcellular location and cell death. *EMBO J*, 18(9):2330–41, May 1999.

[371] Celeste M Nelson, Martijn M Vanduijn, Jamie L Inman, Daniel A Fletcher, and Mina J Bissell. Tissue geometry determines sites of mammary branching morphogenesis in organotypic cultures. *Science*, 314(5797):298–300, Oct 2006.

[372] Kirsten Neubert, Silke Meister, Katrin Moser, Florian Weisel, Damian Maseda, Kerstin Amann, Carsten Wiethe, Thomas H Winkler, Joachim R Kalden, Rudolf A Manz, and Reinhard E Voll. The proteasome inhibitor bortezomib depletes plasma cells and protects mice with lupus-like disease from nephritis. *Nat Med*, 14(7):748–55, Jul 2008.

[373] M E J Newman and M Girvan. Finding and evaluating community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys*, 69(2 Pt 2):026113, Feb 2004.

[374] Duc Thang Nguyên, Sem Kebache, Ali Fazel, Hetty N Wong, Sarah Jenna, Anouk Emadali, Eun-Hye Lee, John J M Bergeron, Randal J Kaufman, Louise Larose, and Eric Chevet. Nck-dependent activation of extracellular signal-regulated kinase-1 and regulation of cell survival during endoplasmic reticulum stress. *Mol Biol Cell*, 15(9):4248–60, Sep 2004.

[375] Min Ni, Hui Zhou, Shiuan Wey, Peter Baumeister, and Amy S Lee. Regulation of perk signaling and leukemic cell survival by a novel cytosolic isoform of the upr regulator grp78/bip. *PLoS One*, 4(8):e6868, 2009.

[376] P Nicotera and S Orrenius. The role of calcium in apoptosis. *Cell Calcium*, 23(2-3):173–80, 1998.

[377] H Nishitoh, M Saitoh, Y Mochida, K Takeda, H Nakano, M Rothe, K Miyazono, and H Ichijo. Ask1 is essential for jnk/sapk activation by traf2. *Mol Cell*, 2(3):389–95, Sep 1998.

[378] Hideki Nishitoh, Hisae Kadowaki, Atsushi Nagai, Takeshi Maruyama, Takanori Yokota, Hisashi Fukutomi, Takuya Noguchi, Atsushi Matsuzawa, Kohsuke Takeda, and Hidenori Ichijo. Als-linked mutant sod1 induces er stress- and ask1-dependent motor neuron death by targeting derlin-1. *Genes Dev*, 22:1451–64, Jun 2008.

[379] M Niwa, C Sidrauski, R J Kaufman, and P Walter. A role for presenilin-1 in nuclear accumulation of ire1 fragments and induction of the mammalian unfolded protein response. *Cell*, 99:691–702, Dec 1999.

[380] T Noda and Y Ohsumi. Tor, a phosphatidylinositol kinase homologue, controls autophagy in yeast. *J Biol Chem*, 273:3963–6, Feb 1998.

[381] N Normanno, C Bianco, L Strizzi, M Mancino, M R Maiello, A De Luca, F Caponigro, and D S Salomon. The erbb receptors and their ligands in cancer: an overview. *Curr Drug Targets*, 6(3):243–57, May 2005.

[382] Nicola Normanno, Antonella De Luca, Caterina Bianco, Luigi Strizzi, Mario Mancino, Monica R Maiello, Adele Carotenuto, Gianfranco De Feo, Francesco Caponigro, and David S Salomon. Epidermal growth factor receptor (egfr) signaling in cancer. *Gene*, 366(1):2–16, Jan 2006.

[383] Nicola Normanno, Antonella De Luca, Monica R Maiello, Mario Mancino, Antonio D'Antonio, Marcella Macaluso, Francesco Caponigro, and Antonio Giordano. Epidermal growth factor receptor (egfr) tyrosine kinase

inhibitors in breast cancer: current status and future development. *Front Biosci*, 10:2611–7, 2005.

[384] I Novoa, H Zeng, H P Harding, and D Ron. Feedback inhibition of the unfolded protein response by gadd34-mediated dephosphorylation of eif2alpha. *J Cell Biol*, 153:1011–22, May 2001.

[385] M Oft, J Peli, C Rudaz, H Schwarz, H Beug, and E Reichmann. Tgf-beta1 and ha-ras collaborate in modulating the phenotypic plasticity and invasiveness of epithelial tumor cells. *Genes Dev*, 10(19):2462–77, Oct 1996.

[386] Maiko Ogata, Shin-ichiro Hino, Atsushi Saito, Keisuke Morikawa, Shinichi Kondo, Soshi Kanemoto, Tomohiko Murakami, Manabu Taniguchi, Ichiro Tanii, Kazuya Yoshinaga, Sadao Shiosaka, James A Hammarback, Fumihiko Urano, and Kazunori Imaizumi. Autophagy is activated for cell survival after endoplasmic reticulum stress. *Mol Cell Biol*, 26:9220–31, Dec 2006.

[387] Tomoshi Ohya, Tetsuji Hayashi, Eriko Kiyama, Hiroko Nishii, Hideo Miki, Kaoru Kobayashi, Kohsuke Honda, Takeshi Omasa, and Hisao Ohtake. Improved production of recombinant human antithrombin iii in chinese hamster ovary cells by atf4 overexpression. *Biotechnol Bioeng*, 100(2):317–24, Jun 2008.

[388] Ken-ichiro Okada, Tetsuo Minamino, Yoshitane Tsukamoto, Yulin Liao, Osamu Tsukamoto, Seiji Takashima, Akio Hirata, Masashi Fujita, Yoko Nagamachi, Takeshi Nakatani, Chikao Yutani, Kentaro Ozawa, Satoshi Ogawa, Hitonobu Tomoike, Masatsugu Hori, and Masafumi Kitakaze. Prolonged endoplasmic reticulum stress in hypertrophic and failing heart after aortic constriction: possible contribution of endoplasmic reticulum stress to cardiac myocyte apoptosis. *Circulation*, 110(6):705–12, Aug 2004.

[389] K Okamura, Y Kimata, H Higashio, A Tsuru, and K Kohno. Dissociation of kar2p/bip from an er sensory molecule, ire1p, triggers the unfolded protein response in yeast. *Biochem Biophys Res Commun*, 279(2):445–50, Dec 2000.

[390] M A Olayioye, R M Neve, H A Lane, and N E Hynes. The erbb signaling network: receptor heterodimerization in development and cancer. *EMBO J*, 19(13):3159–67, Jul 2000.

[391] Silvia Olivari, Carmela Galli, Heli Alanen, Lloyd Ruddock, and Maurizio Molinari. A novel stress-induced edem variant regulating endoplasmic

reticulum-associated glycoprotein degradation. *J Biol Chem*, 280(4):2424–8, Jan 2005.

[392] J D Oliver, H L Roderick, D H Llewellyn, and S High. Erp57 functions as a subunit of specific complexes formed with the er lectins calreticulin and calnexin. *Mol Biol Cell*, 10(8):2573–82, Aug 1999.

[393] J D Oliver, F J van der Wal, N J Bulleid, and S High. Interaction of the thiol-dependent reductase erp57 with nascent glycoproteins. *Science*, 275(5296):86–8, Jan 1997.

[394] Takeshi Omasa, Takashi Takami, Tomoshi Ohya, Eriko Kiyama, Tetsuji Hayashi, Hiroko Nishii, Hideo Miki, Kaoru Kobayashi, Kohsuke Honda, and Hisao Ohtake. Overexpression of gadd34 enhances production of recombinant human antithrombin iii in chinese hamster ovary cells. *J Biosci Bioeng*, 106(6):568–73, Dec 2008.

[395] S. Oyadomari and M. Mori. Roles of CHOP/GADD153 in endoplasmic reticulum stress. *Cell Death & Differentiation*, 11(4):381–389, 2003.

[396] S Oyadomari, K Takeda, M Takiguchi, T Gotoh, M Matsumoto, I Wada, S Akira, E Araki, and M Mori. Nitric oxide-induced apoptosis in pancreatic beta cells is mediated by the endoplasmic reticulum stress pathway. *Proc Natl Acad Sci U S A*, 98(19):10845–50, Sep 2001.

[397] Seiichi Oyadomari, Akio Koizumi, Kiyoshi Takeda, Tomomi Gotoh, Shizuo Akira, Eiichi Araki, and Masataka Mori. Targeted disruption of the chop gene delays endoplasmic reticulum stress-mediated diabetes. *J Clin Invest*, 109(4):525–32, Feb 2002.

[398] Seiichi Oyadomari, Chi Yun, Edward A Fisher, Nicola Kreglinger, Gert Kreibich, Miho Oyadomari, Heather P Harding, Alan G Goodman, Hanna Harant, Jennifer L Garrison, Jack Taunton, Michael G Katze, and David Ron. Cotranslocational degradation protects the stressed endoplasmic reticulum from protein overload. *Cell*, 126(4):727–39, Aug 2006.

[399] Umut Ozcan, Qiong Cao, Erkan Yilmaz, Ann-Hwee Lee, Neal N Iwakoshi, Esra Ozdelen, Gürol Tuncman, Cem Görgün, Laurie H Glimcher, and Gökhan S Hotamisligil. Endoplasmic reticulum stress links obesity, insulin action, and type 2 diabetes. *Science*, 306(5695):457–61, Oct 2004.

[400] Alok S Pachori, Luis G Melo, Melanie L Hart, Nicholas Noiseux, Lunan Zhang, Fulvio Morello, Scott D Solomon, Gregory L Stahl, Richard E Pratt, and Victor J Dzau. Hypoxia-regulated therapeutic gene as a preemptive treatment strategy against ischemia/reperfusion tissue injury. *Proc Natl Acad Sci U S A*, 101(33):12282–7, Aug 2004.

[401] T.N Palmer, G.J Shutts, R Hagedorn, F.J Doblas-Reyes, Y Jung, and M. Leutbecher. Representing model uncertainty in weather and climate prediction. *Ann Rev Earth and Planetary Sci*, 33:163–193, 2005.

[402] Yan-Xia Pan, Li Lin, An-Jing Ren, Xiu-Jie Pan, Hong Chen, Chao-Shu Tang, and Wen-Jun Yuan. Hsp70 and grp78 induced by endothelin-1 pretreatment enhance tolerance to hypoxia in cultured neonatal rat cardiomyocytes. *J Cardiovasc Pharmacol*, 44 Suppl 1:S117–20, Nov 2004.

[403] Feroz R Papa, Chao Zhang, Kevan Shokat, and Peter Walter. Bypassing a kinase activity with an atp-competitive drug. *Science*, 302(5650):1533–7, Nov 2003.

[404] Jason A Papin, Tony Hunter, Bernhard O Palsson, and Shankar Subramaniam. Reconstruction of cellular signalling networks and analysis of their properties. *Nat Rev Mol Cell Biol*, 6(2):99–111, Feb 2005.

[405] Jason A Papin and Bernhard O Palsson. The jak-stat signaling network in the human b-cell: an extreme signaling pathway analysis. *Biophys J*, 87(1):37–46, Jul 2004.

[406] Jason A Papin and Bernhard O Palsson. Topological analysis of mass-balanced signaling networks: a framework to obtain network properties including crosstalk. *J Theor Biol*, 227(2):283–97, Mar 2004.

[407] Jason A Papin, Nathan D Price, and Bernhard Ø Palsson. Extreme pathway lengths and reaction participation in genome-scale metabolic networks. *Genome Res*, 12(12):1889–900, Dec 2002.

[408] Jason A Papin, Jennifer L Reed, and Bernhard O Palsson. Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. *Trends Biochem Sci*, 29(12):641–7, Dec 2004.

[409] Alexander A Parikh, Fan Fan, Wen Biao Liu, Syed A Ahmad, Oliver Stoeltzing, Niels Reinmuth, Diane Bielenberg, Corazon D Bucana, Michael Klagsbrun, and Lee M Ellis. Neuropilin-1 in human colon cancer:

expression, regulation, and role in induction of angiogenesis. *Am J Pathol*, 164(6):2139–51, Jun 2004.

[410] Hae-Ryong Park, Akihiro Tomida, Shigeo Sato, Yoshinori Tsukumo, Jisoo Yun, Takao Yamori, Yoichi Hayakawa, Takashi Tsuruo, and Kazuo Shinya. Effect on tumor cells of blocking survival response to glucose deprivation. *J Natl Cancer Inst*, 96(17):1300–10, Sep 2004.

[411] A J Parodi. Protein glucosylation and its role in protein folding. *Annu Rev Biochem*, 69:69–93, 2000.

[412] W Paschen, C Gissel, T Linden, S Althausen, and J Doutheil. Activation of gadd153 expression through transient cerebral ischemia: evidence that ischemia causes endoplasmic reticulum dysfunction. *Brain Res Mol Brain Res*, 60(1):115–22, Sep 1998.

[413] K. Paz, R. Hemi, D. LeRoith, A. Karasik, E. Elhanany, H. Kanety, and Y. Zick. A molecular basis for insulin resistance. Elevated serine/threonine phosphorylation of IRS-1 and IRS-2 inhibits their binding to the juxtamembrane region of the insulin receptor and impairs their ability to undergo insulin-induced tyrosine phosphorylation. *J Biol Chem*, 272(47):29911–29918, Nov 1997.

[414] K. Paz, Y. F. Liu, H. Shorer, R. Hemi, D. LeRoith, M. Quan, H. Kanety, R. Seger, and Y. Zick. Phosphorylation of insulin receptor substrate-1 (IRS-1) by protein kinase B positively regulates IRS-1 function. *J Biol Chem*, 274(40):28816–28822, Oct 1999.

[415] Dana Pe'er. Bayesian network analysis of signaling networks: a primer. *Sci STKE*, 2005(281):pl4, Apr 2005.

[416] Mark D Pegram and David M Reese. Combined biological therapy of breast cancer using monoclonal antibodies directed against her2/neu protein and vascular endothelial growth factor. *Semin Oncol*, 29(3 Suppl 11):29–37, Jun 2002.

[417] Selma Pennacchietti, Paolo Michieli, Maria Galluzzo, Massimiliano Mazzone, Silvia Giordano, and Paolo M Comoglio. Hypoxia promotes invasive growth by transcriptional activation of the met protooncogene. *Cancer Cell*, 3(4):347–61, Apr 2003.

[418] Holger Perfahl, Helen M Byrne, Tingan Chen, Veronica Estrella, Tomás Alarcón, Alexei Lapin, Robert A Gatenby, Robert J Gillies, Mark C Lloyd,

Philip K Maini, Matthias Reuss, and Markus R Owen. Multiscale modelling of vascular tumour growth in 3d: the roles of domain size and boundary conditions. *PLoS One*, 6(4):e14790, 2011.

[419] C M Perou, T Sørlie, M B Eisen, M van de Rijn, S S Jeffrey, C A Rees, J R Pollack, D T Ross, H Johnsen, L A Akslen, O Fluge, A Pergamenschikov, C Williams, S X Zhu, P E Lønning, A L Børresen-Dale, P O Brown, and D Botstein. Molecular portraits of human breast tumours. *Nature*, 406(6797):747–52, Aug 2000.

[420] M Petersen, E Pardali, G van der Horst, H Cheung, C van den Hoogen, G van der Pluijm, and P Ten Dijke. Smad2 and smad3 have opposing roles in breast cancer bone metastasis by differentially affecting tumor angiogenesis. *Oncogene*, 29(9):1351–61, Mar 2010.

[421] A M Petit, J Rak, M C Hung, P Rockwell, N Goldstein, B Fendly, and R S Kerbel. Neutralizing antibodies against epidermal growth factor and erbb-2/neu receptor tyrosine kinases down-regulate vascular endothelial growth factor production by tumor cells in vitro and in vivo: angiogenic implications for signal transduction therapy of solid tumors. *Am J Pathol*, 151(6):1523–30, Dec 1997.

[422] Kseniya Petrova, Seiichi Oyadomari, Linda M Hendershot, and David Ron. Regulated association of misfolded endoplasmic reticulum lumenal proteins with p58/dnajc3. *EMBO J*, 27(21):2862–72, Nov 2008.

[423] Leonard Petrucelli, Casey O'Farrell, Paul J Lockhart, Melisa Baptista, Kathryn Kehoe, Liselot Vink, Peter Choi, Benjamin Wolozin, Matthew Farrer, John Hardy, and Mark R Cookson. Parkin protects against the toxicity associated with mutant alpha-synuclein: proteasome dysfunction selectively affects catecholaminergic neurons. *Neuron*, 36(6):1007–19, Dec 2002.

[424] D D Pittman, K N Tomkinson, and R J Kaufman. Post-translational requirements for functional factor v and factor viii secretion in mammalian cells. *J Biol Chem*, 269(25):17329–37, Jun 1994.

[425] Rexxi D Prasasya, Dan Tian, and Pamela K Kreeger. Analysis of cancer signaling networks by systems biology to develop therapies. *Semin Cancer Biol*, 21(3):200–6, Jun 2011.

[426] Aleix Prat and Charles M Perou. Deconstructing the molecular portraits of breast cancer. *Mol Oncol*, 5(1):5–23, Feb 2011.

[427] D J Price, T Miralem, S Jiang, R Steinberg, and H Avraham. Role of vascular endothelial growth factor in the stimulation of cellular invasion and signaling of breast cancer cells. *Cell Growth Differ*, 12(3):129–35, Mar 2001.

[428] Nathan D Price, Jennifer L Reed, Jason A Papin, Iman Famili, and Bernhard O Palsson. Analysis of metabolic capabilities using singular value decomposition of extreme pathway matrices. *Biophys J*, 84(2 Pt 1):794–804, Feb 2003.

[429] Christopher W Pugh and Peter J Ratcliffe. Regulation of angiogenesis by hypoxia: role of the hif system. *Nat Med*, 9(6):677–84, Jun 2003.

[430] Girish V Putcha, Siyuan Le, Stephan Frank, Cagri G Besirli, Kim Clark, Boyang Chu, Shari Alix, Richard J Youle, Art LaMarche, Anna C Maroney, and Eugene M Johnson, Jr. Jnk-mediated bim phosphorylation potentiates bax-dependent apoptosis. *Neuron*, 38(6):899–914, Jun 2003.

[431] H Puthalakath, D C Huang, L A O'Reilly, S M King, and A Strasser. The proapoptotic activity of the bcl-2 family member bim is regulated by interaction with the dynein motor complex. *Mol Cell*, 3(3):287–96, Mar 1999.

[432] Hamsa Puthalakath, Lorraine A O'Reilly, Priscilla Gunn, Lily Lee, Priscilla N Kelly, Nicholas D Huntington, Peter D Hughes, Ewa M Michalak, Jennifer McKimm-Breschkin, Noburo Motoyama, Tomomi Gotoh, Shizuo Akira, Philippe Bouillet, and Andreas Strasser. Er stress triggers apoptosis by activating bh3-only protein bim. *Cell*, 129(7):1337–49, Jun 2007.

[433] Peter Pyrko, Adel Kardosh, Weijun Wang, Wenyong Xiong, Axel H Schönthal, and Thomas C Chen. Hiv-1 protease inhibitors nelfinavir and atazanavir induce malignant glioma death by triggering endoplasmic reticulum stress. *Cancer Res*, 67(22):10920–8, Nov 2007.

[434] Amina A Qutub, Gang Liu, Prakash Vempati, and Aleksander S Popel. Integration of angiogenesis modules at multiple scales: from molecular to tissue. *Pac Symp Biocomput*, pages 316–27, 2009.

[435] Amina A Qutub, Feilim Mac Gabhann, Emmanouil D Karagiannis, Prakash Vempati, and Aleksander S Popel. Multiscale models of angiogenesis. *IEEE Eng Med Biol Mag*, 28(2):14–31, 2009.

[436] S Rajagopalan, Y Xu, and M B Brenner. Retention of unassembled compo-

nents of integral membrane proteins by calnexin. *Science*, 263(5145):387–90, Jan 1994.

[437] Aparna C Ranganathan, Lin Zhang, Alejandro P Adam, and Julio A Aguirre-Ghiso. Functional coupling of p38-induced up-regulation of bip and activation of rna-dependent protein kinase-like endoplasmic reticulum kinase to drug resistance of dormant carcinoma cells. *Cancer Res*, 66(3):1702–11, Feb 2006.

[438] R V Rao, E Hermel, S Castro-Obregon, G del Rio, L M Ellerby, H M Ellerby, and D E Bredesen. Coupling endoplasmic reticulum stress to the cell death program. mechanism of caspase activation. *J Biol Chem*, 276(36):33869–74, Sep 2001.

[439] Rammohan V Rao and Dale E Bredesen. Misfolded proteins, endoplasmic reticulum stress and neurodegeneration. *Curr Opin Cell Biol*, 16(6):653–62, Dec 2004.

[440] R.V. Rao, S. Castro-Obregon, H. Frankowski, M. Schuler, V. Stoka, G. del Rio, D.E. Bredesen, and H.M. Ellerby. Coupling Endoplasmic Reticulum Stress to the Cell Death Program AN Apaf-1-INDEPENDENT INTRINSIC PATHWAY. *J. Biol. Chem.*, 277(24):21836–21842, 2002.

[441] A. Rapisarda and G. Melillo. Combination Strategies Targeting Hypoxia Inducible Factor 1 (HIF-1) for Cancer Therapy. *The Tumor Microenvironment*, pages 3–21, 2010.

[442] B. Raught, A. C. Gingras, S. P. Gygi, H. Imataka, S. Morino, A. Gradi, R. Aebersold, and N. Sonenberg. Serum-stimulated, rapamycin-sensitive phosphorylation sites in the eukaryotic translation initiation factor 4GI. *EMBO J*, 19(3):434–444, Feb 2000.

[443] Brian Raught, Franck Peiretti, Anne-Claude Gingras, Mark Livingstone, David Shahbazian, Greg L Mayeur, Roberto D Polakiewicz, Nahum Sonenberg, and John W B Hershey. Phosphorylation of eucaryotic translation initiation factor 4B Ser422 is modulated by S6 kinases. *EMBO J*, 23(8):1761–1769, Apr 2004.

[444] Jennifer F Raven, Dionissios Baltzis, Shuo Wang, Zineb Mounir, Andreas I Papadakis, Hong Qing Gao, and Antonis E Koromilas. Pkr and pkr-like endoplasmic reticulum kinase induce the proteasome-dependent degradation of cyclin d1 via a mechanism requiring eukaryotic initiation factor 2alpha phosphorylation. *J Biol Chem*, 283(6):3097–108, Feb 2008.

[445] R Ravi, B Mookerjee, Z M Bhujwalla, C H Sutter, D Artemov, Q Zeng, L E Dillehay, A Madan, G L Semenza, and A Bedi. Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1alpha. *Genes Dev*, 14(1):34–44, Jan 2000.

[446] Mahboobeh Razmkhah, Mansooreh Jaberipour, Ahmad Hosseini, Akbar Safaei, Behzad Khalatbari, and Abbas Ghaderi. Expression profile of il-8 and growth factors in breast cancer cells and adipose-derived stem cells (ascs) isolated from breast carcinoma. *Cell Immunol*, 265(1):80–5, 2010.

[447] Ramachandra K Reddy, Changhui Mao, Peter Baumeister, Richard C Austin, Randal J Kaufman, and Amy S Lee. Endoplasmic reticulum chaperone protein grp78 protects cells from apoptosis induced by topoisomerase inhibitors: role of atp binding site in suppression of caspase-7 activation. *J Biol Chem*, 278(23):20915–24, Jun 2003.

[448] Jennifer L Reed, Thuy D Vo, Christophe H Schilling, and Bernhard O Palsson. An expanded genome-scale model of escherichia coli k-12 (ijr904 gsm/gpr). *Genome Biol*, 4(9):R54, 2003.

[449] Sami Reijonen, Noora Putkonen, Anne Nørremølle, Dan Lindholm, and Laura Korhonen. Inhibition of endoplasmic reticulum stress counteracts neuronal cell death and protein aggregation caused by n-terminal mutant huntingtin proteins. *Exp Cell Res*, 314(5):950–60, Mar 2008.

[450] J. M. Ricort, J. F. Tanti, E. Van Obberghen, and Y. Le Marchand-Brustel. Cross-talk between the platelet-derived growth factor and the insulin signaling pathways in 3T3-L1 adipocytes. *J Biol Chem*, 272(32):19814–19818, Aug 1997.

[451] Jordi Rius, Monica Guma, Christian Schachtrup, Katerina Akassoglou, Annelies S Zinkernagel, Victor Nizet, Randall S Johnson, Gabriel G Haddad, and Michael Karin. Nf-kappab links innate immunity to the hypoxic response through transcriptional regulation of hif-1alpha. *Nature*, 453(7196):807–11, Jun 2008.

[452] P. Rodriguez-Viciana, P. H. Warne, R. Dhand, B. Vanhaesebroeck, I. Gout, M. J. Fry, M. D. Waterfield, and J. Downward. Phosphatidylinositol-3-OH kinase as a direct target of Ras. *Nature*, 370(6490):527–532, Aug 1994.

[453] E K Rofstad and E F Halsør. Hypoxia-associated spontaneous pulmonary metastasis in human melanoma xenografts: involvement of microvas-

cular hot spots induced in hypoxic foci by interleukin 8. *Br J Cancer*, 86(2):301–8, Jan 2002.

[454] Lorenzo Romero-Ramirez, Hongbin Cao, Daniel Nelson, Ester Hammond, Ann-Hwee Lee, Hiderou Yoshida, Kazutoshi Mori, Laurie H Glimcher, Nicholas C Denko, Amato J Giaccia, Quynh-Thu Le, and Albert C Koong. Xbp1 is essential for survival under hypoxic conditions and is required for tumor growth. *Cancer Res*, 64(17):5943–7, Sep 2004.

[455] D Ron and J F Habener. Chop, a novel developmentally regulated nuclear protein that dimerizes with transcription factors c/ebp and lap and functions as a dominant-negative inhibitor of gene transcription. *Genes Dev*, 6(3):439–53, Mar 1992.

[456] David Ron. Translational control in the endoplasmic reticulum stress response. *J Clin Invest*, 110(10):1383–8, Nov 2002.

[457] David Ron and Peter Walter. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol*, 8(7):519–29, Jul 2007.

[458] Robert Roskoski, Jr. Vascular endothelial growth factor (vegf) signaling in tumor progression. *Crit Rev Oncol Hematol*, 62(3):179–213, Jun 2007.

[459] Jean-François Rual, Kavitha Venkatesan, Tong Hao, Tomoko Hirozane-Kishikawa, Amélie Dricot, Ning Li, Gabriel F Berriz, Francis D Gibbons, Matija Dreze, Nono Ayivi-Guedehoussou, Niels Klitgord, Christophe Simon, Mike Boxem, Stuart Milstein, Jennifer Rosenberg, Debra S Goldberg, Lan V Zhang, Sharyl L Wong, Giovanni Franklin, Siming Li, Joanna S Albala, Janghoo Lim, Carlene Fraughton, Estelle Llamosas, Sebiha Cevik, Camille Bex, Philippe Lamesch, Robert S Sikorski, Jean Vandenhaute, Huda Y Zoghbi, Alex Smolyar, Stephanie Bosak, Reynaldo Sequerra, Lynn Doucette-Stamm, Michael E Cusick, David E Hill, Frederick P Roth, and Marc Vidal. Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, 437(7062):1173–8, Oct 2005.

[460] Derek Ruths, Melissa Muller, Jen-Te Tseng, Luay Nakhleh, and Prahlad T Ram. The signaling petri net-based simulator: a non-parametric strategy for characterizing the dynamics of cell-specific signaling networks. *PLoS Comput Biol*, 4(2):e1000005, Feb 2008.

[461] Elizabeth J Ryu, Heather P Harding, James M Angelastro, Ottavio V Vitolo, David Ron, and Lloyd A Greene. Endoplasmic reticulum stress and

the unfolded protein response in cellular models of parkinson's disease. *J Neurosci*, 22(24):10690–8, Dec 2002.

[462] Deepali Sachdev and Douglas Yee. Disrupting insulin-like growth factor signaling as a potential cancer therapy. *Mol Cancer Ther*, 6(1):1–12, Jan 2007.

[463] Andrea Sackmann, Monika Heiner, and Ina Koch. Application of petri net based analysis techniques to signal transduction pathways. *BMC Bioinformatics*, 7:482, 2006.

[464] Maya Saleh, John C Mathison, Melissa K Wolinski, Steve J Bensinger, Patrick Fitzgerald, Nathalie Droin, Richard J Ulevitch, Douglas R Green, and Donald W Nicholson. Enhanced bacterial clearance and sepsis resistance in caspase-12-deficient mice. *Nature*, 440(7087):1064–8, Apr 2006.

[465] Leonardo Salmena, Arkaitz Carracedo, and Pier Paolo Pandolfi. Tenets of PTEN tumor suppression. *Cell*, 133(3):403–414, May 2008.

[466] A. R. Saltiel and C. R. Kahn. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature*, 414(6865):799–806, Dec 2001.

[467] Alan R Saltiel and Jeffrey E Pessin. Insulin signaling pathways in time and space. *Trends Cell Biol*, 12(2):65–71, Feb 2002.

[468] Yardena Samuels, Zhenghe Wang, Alberto Bardelli, Natalie Silliman, Janine Ptak, Steve Szabo, Hai Yan, Adi Gazdar, Steven M Powell, Gregory J Riggins, James K V Willson, Sanford Markowitz, Kenneth W Kinzler, Bert Vogelstein, and Victor E Velculescu. High frequency of mutations of the PIK3CA gene in human cancers. *Science*, 304(5670):554, Apr 2004.

[469] Yasemin Sancak, Carson C Thoreen, Timothy R Peterson, Robert A Lindquist, Seong A Kang, Eric Spooner, Steven A Carr, and David M Sabatini. PRAS40 is an insulin-regulated inhibitor of the mTORC1 protein kinase. *Mol Cell*, 25(6):903–915, Mar 2007.

[470] Ignacio E Sánchez, Manuel Morillas, Eva Zobeley, Thomas Kiefhaber, and Rudi Glockshuber. Fast folding of the two-domain semliki forest virus capsid protein explains co-translational proteolytic activity. *J Mol Biol*, 338(1):159–67, Apr 2004.

[471] Werner Sandmann. Discrete-time stochastic modeling and simulation of biochemical networks. *Comput Biol Chem*, 32(4):292–7, Aug 2008.

[472] D. D. Sarbassov, David A Guertin, Siraj M Ali, and David M Sabatini. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science*, 307(5712):1098–1101, Feb 2005.

[473] Leslie J Saucedo, Xinsheng Gao, Dominic A Chiarelli, Ling Li, Duoija Pan, and Bruce A Edgar. Rheb promotes cell growth as a component of the insulin/TOR signalling network. *Nat Cell Biol*, 5(6):566–571, Jun 2003.

[474] Gert C Scheper and Christopher G Proud. Does phosphorylation of the cap-binding protein eIF4E play a role in translation initiation? *Eur J Biochem*, 269(22):5350–5359, Nov 2002.

[475] D Scheuner, B Song, E McEwen, C Liu, R Laybutt, P Gillespie, T Saunders, S Bonner-Weir, and R J Kaufman. Translational control is required for the unfolded protein response and in vivo glucose homeostasis. *Mol Cell*, 7(6):1165–76, Jun 2001.

[476] F. Schmitt. Her2+ breast cancer: how to evaluate? *Adv Ther*, 26 Suppl 1:S1–S8, Jul 2009.

[477] R.J. Schneider, N. Sonenberg, M. Mathews, N. Sonenberg, and J. Hershey. Translational control in cancer development and progression. *Translational control in biology and medicine*, pages 401–431, 2007.

[478] Santiago Schnell. A model of the unfolded protein response: pancreatic beta-cell as a case study. *Cell Physiol Biochem*, 23(4-6):233–44, 2009.

[479] Birgit Schoeberl, Claudia Eichler-Jonsson, Ernst Dieter Gilles, and Gertraud Müller. Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nat Biotechnol*, 20(4):370–5, Apr 2002.

[480] J D Schrag, J J Bergeron, Y Li, S Borisova, M Hahn, D Y Thomas, and M Cygler. The structure of calnexin, an er chaperone involved in quality control of protein folding. *Mol Cell*, 8(3):633–44, Sep 2001.

[481] Martin Schröder and Randal J Kaufman. The mammalian unfolded protein response. *Annu Rev Biochem*, 74:739–89, 2005.

[482] S Schuster, D A Fell, and T Dandekar. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat Biotechnol*, 18(3):326–32, Mar 2000.

[483] D J Selkoe. Alzheimer's disease: genes, proteins, and therapy. *Physiol Rev*, 81(2):741–66, Apr 2001.

[484] Gregg L Semenza. Hif-1 and tumor progression: pathophysiology and therapeutics. *Trends Mol Med*, 8(4 Suppl):S62–7, 2002.

[485] S E Shackney. A computer model for tumor growth and chemotherapy, and its application to l1210 leukemia treated with cytosine arabinoside (nsc-63878). *Cancer Chemother Rep*, 54(6):399–429, Dec 1970.

[486] C E Shamu and P Walter. Oligomerization and phosphorylation of the ire1p kinase during intracellular signaling from the endoplasmic reticulum to the nucleus. *EMBO J*, 15(12):3028–39, Jun 1996.

[487] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13(11):2498–504, Nov 2003.

[488] Li Shao, Xiujun Sun, Li Xu, L Trevor Young, and Jun-Feng Wang. Mood stabilizing drug lithium increases expression of endoplasmic reticulum stress proteins in primary cultured rat cerebral cortical cells. *Life Sci*, 78(12):1317–23, Feb 2006.

[489] Jingshi Shen, Xi Chen, Linda Hendershot, and Ron Prywes. Er stress regulation of atf6 localization by dissociation of bip/grp78 binding and unmasking of golgi localization signals. *Dev Cell*, 3:99–111, Jul 2002.

[490] Jingshi Shen and Ron Prywes. Dependence of site-2 protease cleavage of atf6 on prior site-1 protease digestion is determined by the size of the luminal domain of atf6. *J Biol Chem*, 279:43046–51, Oct 2004.

[491] X Shen, R E Ellis, K Lee, C Y Liu, K Yang, A Solomon, H Yoshida, R Morimoto, D M Kurnit, K Mori, and R J Kaufman. Complementary signaling pathways regulate the unfolded protein response and are required for c. elegans development. *Cell*, 107:893–903, Dec 2001.

[492] Xiaohua Shen, Kezhong Zhang, and Randal J Kaufman. The unfolded

protein response–a stress signaling pathway of the endoplasmic reticulum. *J Chem Neuroanat*, 28(1-2):79–92, Sep 2004.

[493] Y Shi, J An, J Liang, S E Hayes, G E Sandusky, L E Stramm, and N N Yang. Characterization of a mutant pancreatic eif-2alpha kinase, pek, and colocalization with somatostatin in islet delta cells. *J Biol Chem*, 274(9):5723–30, Feb 1999.

[494] Y Shi, K M Vattem, R Sood, J An, J Liang, L Stramm, and R C Wek. Identification and characterization of pancreatic eukaryotic initiation factor 2 alpha-subunit kinase, pek, involved in translational control. *Mol Cell Biol*, 18(12):7499–509, Dec 1998.

[495] Andrew J Shih, Jeremy Purvis, and Ravi Radhakrishnan. Molecular systems biology of erbb1 signaling: bridging the gap through multiscale modeling and high-performance computing. *Mol Biosyst*, 4(12):1151–9, Dec 2008.

[496] Kaori Shintani-Ishida, Makoto Nakajima, Koichi Uemura, and Ken-ichi Yoshida. Ischemic preconditioning protects cardiomyocytes against ischemic injury by inducing grp78. *Biochem Biophys Res Commun*, 345:1600–5, Jul 2006.

[497] Che-Chi Shu, Anushree Chatterjee, Gary Dunny, Wei-Shou Hu, and Doraiswami Ramkrishna. Bistability versus bimodal distributions in gene regulatory processes from population balance. *PLoS Comput Biol*, 7(8):e1002140, Aug 2011.

[498] Masahiro Shuda, Nobuo Kondoh, Nobuo Imazeki, Kenji Tanaka, Tetsuya Okada, Kazutoshi Mori, Akiyuki Hada, Masaaki Arai, Toru Wakatsuki, Osamu Matsubara, Naoki Yamamoto, and Mikio Yamamoto. Activation of the atf6, xbp1 and grp78 genes in human hepatocellular carcinoma: a possible involvement of the er stress pathway in hepatocarcinogenesis. *J Hepatol*, 38:605–14, May 2003.

[499] C Sidrauski and P Walter. The transmembrane kinase ire1p is a site-specific endonuclease that initiates mrna splicing in the unfolded protein response. *Cell*, 90:1031–9, Sep 1997.

[500] Ariosto Silva, Alexander R A Anderson, and Robert Gatenby. A multiscale model of the bone marrow and hematopoiesis. *Math Biosci Eng*, 8(2):643–58, Apr 2011.

[501] E. Y. Skolnik, A. Batzer, N. Li, C. H. Lee, E. Lowenstein, M. Mohammadi, B. Margolis, and J. Schlessinger. The function of GRB2 in linking the insulin receptor to Ras signaling pathways. *Science*, 260(5116):1953–1955, Jun 1993.

[502] E. Y. Skolnik, C. H. Lee, A. Batzer, L. M. Vicentini, M. Zhou, R. Daly, M. J. Myers, J. M. Backer, A. Ullrich, and M. F. White. The SH2/SH3 domain-containing protein GRB2 interacts with tyrosine-phosphorylated IRS1 and Shc: implications for insulin control of ras signalling. *EMBO J*, 12(5):1929–1936, May 1993.

[503] Noam Slonim, Olivier Elemento, and Saeed Tavazoie. Ab initio genotype-phenotype association reveals intrinsic modularity in genetic networks. *Mol Syst Biol*, 2:2006.0005, 2006.

[504] Sang Ok Song, Anirikh Chakrabarti, and Jeffrey D Varner. Ensembles of signal transduction models using pareto optimal ensemble techniques (poets). *Biotechnol J*, 5(7):768–80, Jul 2010.

[505] Sang Ok Song and Jeffrey Varner. Modeling and analysis of the molecular basis of pain in sensory neurons. *PLoS One*, 4(9):e6758, 2009.

[506] Claudio Soto. Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat Rev Neurosci*, 4(1):49–60, Jan 2003.

[507] M Sousa and A J Parodi. The molecular basis for the recognition of misfolded glycoproteins by the udp-glc:glycoprotein glucosyltransferase. *EMBO J*, 14(17):4196–203, Sep 1995.

[508] Jaroslav Stark, Robin Callard, and Michael Hubank. From the top down: towards a predictive biology of signalling networks. *Trends Biotechnol*, 21(7):290–3, Jul 2003.

[509] Marianne O Stefanini, Amina A Qutub, Feilim Mac Gabhann, and Aleksander S Popel. Computational models of vegf-associated angiogenic processes in cancer. *Math Med Biol*, Sep 2011.

[510] J Stelling, E D Gilles, and F J Doyle. Robustness properties of circadian clock architectures. *Proc Natl Acad Sci U S A*, 101(36):13210–13215, 2004.

[511] Jörg Stelling, Steffen Klamt, Katja Bettenbrock, Stefan Schuster, and

Ernst Dieter Gilles. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420(6912):190–3, Nov 2002.

[512] Edward C Stites, Paul C Trampont, Zhong Ma, and Kodi S Ravichandran. Network analysis of oncogenic ras activation in cancer. *Science*, 318(5849):463–467, Oct 2007.

[513] H. Sun, R. Lesche, D. M. Li, J. Liliental, H. Zhang, J. Gao, N. Gavrilova, B. Mueller, X. Liu, and H. Wu. PTEN modulates cell cycle progression and cell survival by regulating phosphatidylinositol 3,4,5,-trisphosphate and Akt/protein kinase B signaling pathway. *Proc Natl Acad Sci U S A*, 96(11):6199–6204, May 1999.

[514] Tao Sun, Salem Adra, Rod Smallwood, Mike Holcombe, and Sheila MacNeil. Exploring hypotheses of the actions of tgf-beta1 in epidermal wound healing using a 3d computational multiscale model of the human epidermis. *PLoS One*, 4(12):e8515, 2009.

[515] Eva Szegezdi, Susan E Logue, Adrienne M Gorman, and Afshin Samali. Mediators of endoplasmic reticulum stress-induced apoptosis. *EMBO Rep*, 7(9):880–5, Sep 2006.

[516] S Tajiri, S Oyadomari, S Yano, M Morioka, T Gotoh, J I Hamada, Y Ushio, and M Mori. Ischemia-induced neuronal cell death is mediated by the endoplasmic reticulum stress pathway involving chop. *Cell Death Differ*, 11(4):403–15, Apr 2004.

[517] P P Tak and G S Firestein. Nf-kappab: a key role in inflammatory diseases. *J Clin Invest*, 107(1):7–11, Jan 2001.

[518] Ryosuke Takahashi, Yuzuru Imai, Nobutaka Hattori, and Yoshikuni Mizuno. Parkin and endoplasmic reticulum stress. *Ann N Y Acad Sci*, 991:101–6, Jun 2003.

[519] Yoko Takahashi, Fabiana C Morales, Erica L Kreimann, and Maria-Magdalena Georgescu. PTEN tumor suppressor associates with NHERF proteins to attenuate PDGF receptor signaling. *EMBO J*, 25(4):910–920, Feb 2006.

[520] M Takeichi. Cadherin cell adhesion receptors as a morphogenetic regulator. *Science*, 251(5000):1451–5, Mar 1991.

[521] M Takeichi. Cadherins in cancer: implications for invasion and metastasis. *Curr Opin Cell Biol*, 5(5):806–11, Oct 1993.

[522] C Talcott, S Eker, M Knapp, P Lincoln, and K Laderoute. Pathway logic modeling of protein functional domains in signal transduction. *Pac Symp Biocomput*, pages 568–80, 2004.

[523] Ern Yu Tan, Leticia Campo, Cheng Han, Helen Turley, Francesco Pezzella, Kevin C Gatter, Adrian L Harris, and Stephen B Fox. Bnip3 as a progression marker in primary human breast cancer; opposing functions in in situ versus invasive cancer. *Clin Cancer Res*, 13(2 Pt 1):467–74, Jan 2007.

[524] S Tanaka, T Uehara, and Y Nomura. Up-regulation of protein-disulfide isomerase in response to hypoxia/brain ischemia and its protective effect against apoptotic cell death. *J Biol Chem*, 275(14):10388–93, Apr 2000.

[525] Cullen M Taniguchi, Brice Emanuelli, and C. Ronald Kahn. Critical nodes in signalling pathways: insights into insulin action. *Nat Rev Mol Cell Biol*, 7(2):85–96, Feb 2006.

[526] Ryan Tasseff, Satyaprakash Nayak, Saniya Salim, Poorvi Kaushik, Noreen Rizvi, and Jeffrey D Varner. Analysis of the molecular networks in androgen dependent and independent prostate cancer revealed fragile and robust subsystems. *PLoS One*, 5(1):e8864, 2010.

[527] Ryan Tasseff, Satyaprakash Nayak, Sang Ok Song, Andrew Yen, and Jeffrey D Varner. Modeling and analysis of retinoic acid induced differentiation of uncommitted precursor cells. *Integr Biol (Camb)*, 3(5):578–591, May 2011.

[528] S Tavazoie, J D Hughes, M J Campbell, R J Cho, and G M Church. Systematic determination of genetic network architecture. *Nat Genet*, 22(3):281–5, Jul 1999.

[529] Ian W Taylor, Rune Linding, David Warde-Farley, Yongmei Liu, Catia Pesquita, Daniel Faria, Shelley Bull, Tony Pawson, Quaid Morris, and Jeffrey L Wrana. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol*, 27(2):199–204, Feb 2009.

[530] Jesper Tegner, M K Stephen Yeung, Jeff Hasty, and James J Collins. Reverse engineering gene networks: integrating genetic perturbations with dynamical modeling. *Proc Natl Acad Sci U S A*, 100(10):5944–9, May 2003.

[531] Shannon E Telesco, Andrew J Shih, Fei Jia, and Ravi Radhakrishnan. A multiscale modeling approach to investigate molecular mechanisms of pseudokinase activation and drug resistance in the her3/erbb3 receptor tyrosine kinase signaling network. *Mol Biosyst*, 7(6):2066–80, Jun 2011.

[532] Faraj Terro, Christian Czech, Françoise Esclaire, Wassim Elyaman, Catherine Yardin, Marie-Claire Baclet, Nathalie Touchet, Günter Tremp, Laurent Pradier, and Jacques Hugon. Neurons overexpressing mutant presenilin-1 are more sensitive to apoptosis induced by endoplasmic reticulum-golgi stress. *J Neurosci Res*, 69(4):530–9, Aug 2002.

[533] Jean Paul Thiery. Epithelial-mesenchymal transitions in development and pathologies. *Curr Opin Cell Biol*, 15(6):740–6, Dec 2003.

[534] Jean Paul Thiery and Jonathan P Sleeman. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol*, 7(2):131–42, Feb 2006.

[535] Mahesh Thirunavukkarasu, Bela Juhasz, Lijun Zhan, Venugopal P Menon, Arpad Tosaki, Hajime Otani, and Nilanjana Maulik. Vegfr1 (flt-1+/-) gene knockout leads to the disruption of vegf-mediated signaling through the nitric oxide/heme oxygenase pathway in ischemic preconditioned myocardium. *Free Radic Biol Med*, 42(10):1487–95, May 2007.

[536] Michael J Thomenius and Clark W Distelhorst. Bcl-2 on the endoplasmic reticulum: protecting the mitochondria from a distance. *J Cell Sci*, 116(Pt 22):4493–9, Nov 2003.

[537] Donna J Thuerauf, Marie Marcinko, Peter J Belmont, and Christopher C Glembotski. Effects of the isoform-specific characteristics of atf6 alpha and atf6 beta on endoplasmic reticulum stress response gene expression and cell viability. *J Biol Chem*, 282:22865–78, Aug 2007.

[538] Donna J Thuerauf, Marie Marcinko, Natalie Gude, Marta Rubio, Mark A Sussman, and Christopher C Glembotski. Activation of the unfolded protein response in infarcted mouse heart and hypoxic cultured cardiac myocytes. *Circ Res*, 99:275–82, Aug 2006.

[539] Marcel Tigges and Martin Fussenegger. Xbp1-based engineering of secretory capacity enhances the productivity of chinese hamster ovary cells. *Metab Eng*, Apr 2006.

[540] W Tirasophon, A A Welihinda, and R J Kaufman. A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (ire1p) in mammalian cells. *Genes Dev*, 12:1812–24, Jun 1998.

[541] Boaz Tirosh, Neal N Iwakoshi, Laurie H Glimcher, and Hidde L Ploegh. Rapid turnover of unspliced xbp-1 as a factor that modulates the unfolded protein response. *J Biol Chem*, 281(9):5852–60, Mar 2006.

[542] Kei Tobiume, Masao Saitoh, and Hidenori Ichijo. Activation of apoptosis signal-regulating kinase 1 by the stress-induced activating phosphorylation of pre-formed oligomer. *J Cell Physiol*, 191(1):95–104, Apr 2002.

[543] Derrich J Todd, Ann-Hwee Lee, and Laurie H Glimcher. The endoplasmic reticulum stress response in immunity and autoimmunity. *Nat Rev Immunol*, 8(9):663–74, Sep 2008.

[544] Giampaolo Tortora, Fortunato Ciardiello, and Giampietro Gasparini. Combined targeting of egfr-dependent and vegf-dependent pathways: rationale, preclinical studies and clinical applications. *Nat Clin Pract Oncol*, 5(9):521–30, Sep 2008.

[545] E Sergio Trombetta and Armando J Parodi. Quality control and protein folding in the secretory pathway. *Annu Rev Cell Dev Biol*, 19:649–76, 2003.

[546] Billy Tsai, Yihong Ye, and Tom A Rapoport. Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nat Rev Mol Cell Biol*, 3:246–55, Apr 2002.

[547] J R Tyson and C J Stirling. Lhs1 and sil1 provide a lumenal function that is essential for protein translocation into the endoplasmic reticulum. *EMBO J*, 19(23):6440–52, Dec 2000.

[548] Alexandros Tzatsos and Konstantin V Kandror. Nutrients suppress phosphatidylinositol 3-kinase/Akt signaling via raptor-dependent mTOR-mediated insulin receptor substrate 1 phosphorylation. *Mol Cell Biol*, 26(1):63–76, Jan 2006.

[549] Kohjiro Ueki, Tatsuya Kondo, and C. Ronald Kahn. Suppressor of cytokine signaling 1 (SOCS-1) and SOCS-3 cause insulin resistance through inhibition of tyrosine phosphorylation of insulin receptor substrate proteins by discrete mechanisms. *Mol Cell Biol*, 24(12):5434–5446, Jun 2004.

[550] Sung Hee Um, David D'Alessio, and George Thomas. Nutrient overload, insulin resistance, and ribosomal protein S6 kinase 1, S6K1. *Cell Metab*, 3(6):393–402, Jun 2006.

[551] Sung Hee Um, Francesca Frigerio, Mitsuhiro Watanabe, Frédéric Picard, Manel Joaquin, Melanie Sticker, Stefano Fumagalli, Peter R Allegrini, Sara C Kozma, Johan Auwerx, and George Thomas. Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity. *Nature*, 431(7005):200–205, Sep 2004.

[552] Annika Unruh, Anke Ressel, Hamid G Mohamed, Randall S Johnson, Roger Nadrowitz, Eckart Richter, Dörthe M Katschinski, and Roland H Wenger. The hypoxia-inducible factor-1 alpha is a negative factor for tumor therapy. *Oncogene*, 22(21):3213–20, May 2003.

[553] Ursula Unterberger, Romana Höftberger, Ellen Gelpi, Helga Flicker, Herbert Budka, and Till Voigtländer. Endoplasmic reticulum stress features are prominent in alzheimer disease but not in prion diseases in vivo. *J Neuropathol Exp Neurol*, 65:348–57, Apr 2006.

[554] F Urano, X Wang, A Bertolotti, Y Zhang, P Chung, H P Harding, and D Ron. Coupling of stress in the er to activation of jnk protein kinases by transmembrane protein kinase ire1. *Science*, 287:664–6, Jan 2000.

[555] Masuko Ushio-Fukai and Yoshimasa Nakamura. Reactive oxygen species and angiogenesis: Nadph oxidase as target for cancer therapy. *Cancer Lett*, 266(1):37–52, Jul 2008.

[556] Anne Uv, Rafael Cantera, and Christos Samakovlis. Drosophila tracheal morphogenesis: intricate cellular solutions to basic plumbing problems. *Trends Cell Biol*, 13(6):301–9, Jun 2003.

[557] P J van Diest, R J Michalides, L Jannink, P van der Valk, H L Peterse, J S de Jong, C J Meijer, and J P Baak. Cyclin d1 expression in invasive breast cancer. correlations and prognostic value. *Am J Pathol*, 150(2):705–11, Feb 1997.

[558] Rika van Huizen, Jennifer L Martindale, Myriam Gorospe, and Nikki J Holbrook. P58ipk, a novel endoplasmic reticulum stress-inducible protein and potential negative regulator of eif2alpha signaling. *J Biol Chem*, 278(18):15558–64, May 2003.

[559] Tracy Vargo-Gogola and Jeffrey M Rosen. Modelling breast cancer: one size does not fit all. *Nat Rev Cancer*, 7(9):659–72, Sep 2007.

[560] J D Varner. Large-scale prediction of phenotype: concept. *Biotechnol Bioeng*, 69(6):664–78, Sep 2000.

[561] Sharat Jacob Vayttaden and Upinder Singh Bhalla. Developing complex signaling models using genesis/kinetikit. *Sci STKE*, 2004(219):pl4, Feb 2004.

[562] Laura M Vecchiarelli-Federico, David Cervi, Mehran Haeri, Yanmei Li, Andras Nagy, and Yaacov Ben-David. Vascular endothelial growth factor–a positive and negative regulator of tumor growth. *Cancer Res*, 70(3):863–7, Feb 2010.

[563] Scott S Verbridge, Nak Won Choi, Ying Zheng, Daniel J Brooks, Abraham D Stroock, and Claudia Fischbach. Oxygen-controlled three-dimensional cultures to analyze tumor angiogenesis. *Tissue Eng Part A*, 16(7):2133–41, Jul 2010.

[564] Alexei Verkhratsky. Physiology and pathophysiology of the calcium store in the endoplasmic reticulum of neurons. *Physiol Rev*, 85(1):201–79, Jan 2005.

[565] Jose M G Vilar, Ronald Jansen, and Chris Sander. Signal processing in the tgf-beta superfamily ligand-receptor network. *PLoS Comput Biol*, 2(1):e3, Jan 2006.

[566] Priam Villalonga, Silvia Fernández de Mattos, and Anne J Ridley. RhoE inhibits 4E-BP1 phosphorylation and eIF4E function impairing cap-dependent translation. *J Biol Chem*, 284(51):35287–35296, Dec 2009.

[567] Maurizio Vitadello, Daniele Penzo, Valeria Petronilli, Genny Michieli, Selena Gomirato, Roberta Menabò, Fabio Di Lisa, and Luisa Gorza. Overexpression of the stress protein grp94 reduces cardiomyocyte necrosis due to calcium overload and simulated ischemia. *FASEB J*, 17:923–5, May 2003.

[568] Jiri Vohradsky, Pavel Branny, and Charles J Thompson. Comparative analysis of gene expression on mrna and protein level during development of streptomyces cultures by using singular value decomposition. *Proteomics*, 7(21):3853–66, Nov 2007.

[569] Klaus Voss, Monika Heiner, and Ina Koch. Steady state analysis of metabolic pathways using petri nets. *In Silico Biol*, 3(3):367–87, 2003.

[570] R. G. Vries, A. Flynn, J. C. Patel, X. Wang, R. M. Denton, and C. G. Proud. Heat shock increases the association of binding protein-1 with initiation factor 4E. *J Biol Chem*, 272(52):32779–32784, Dec 1997.

[571] H G Wang, N Pathan, I M Ethell, S Krajewski, Y Yamaguchi, F Shibasaki, F McKeon, T Bobo, T F Franke, and J C Reed. Ca2+-induced apoptosis through calcineurin dephosphorylation of bad. *Science*, 284:339–43, Apr 1999.

[572] Lifu Wang, Thurl E Harris, and John C Lawrence. Regulation of proline-rich Akt substrate of 40 kDa (PRAS40) function by mammalian target of rapamycin complex 1 (mTORC1)-mediated phosphorylation. *J Biol Chem*, 283(23):15619–15627, Jun 2008.

[573] X Z Wang, H P Harding, Y Zhang, E M Jolicoeur, M Kuroda, and D Ron. Cloning of mammalian ire1 reveals diversity in the er stress responses. *EMBO J*, 17:5708–17, Oct 1998.

[574] X Z Wang and D Ron. Stress-induced phosphorylation and activation of the transcription factor chop (gadd153) by p38 map kinase. *Science*, 272:1347–9, May 1996.

[575] Y Wang, J Shen, N Arenzana, W Tirasophon, R J Kaufman, and R Prywes. Activation of atf6 and an atf6 dna binding site by the endoplasmic reticulum stress response. *J Biol Chem*, 275:27013–20, Sep 2000.

[576] Zhihui Wang, Christina M Birch, and Thomas S Deisboeck. Cross-scale sensitivity analysis of a non-small cell lung cancer model: linking molecular signaling properties to cellular behavior. *Biosystems*, 92(3):249–58, Jun 2008.

[577] Zhihui Wang, Christina M Birch, Jonathan Sagotsky, and Thomas S Deisboeck. Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model. *Bioinformatics*, 25(18):2389–96, Sep 2009.

[578] Zhihui Wang, Veronika Bordas, and Thomas S Deisboeck. Discovering molecular targets in cancer with multiscale modeling. *Drug Dev Res*, 72(1):45–52, Feb 2011.

[579] F E Ware, A Vassilakos, P A Peterson, M R Jackson, M A Lehrman, and D B Williams. The molecular chaperone calnexin binds glc1man9glcnac2 oligosaccharide as an initial step in recognizing unfolded glycoproteins. *J Biol Chem*, 270(9):4697–704, Mar 1995.

[580] David J J Waugh and Catherine Wilson. The interleukin-8 pathway in cancer. *Clin Cancer Res*, 14(21):6735–41, Nov 2008.

[581] M C Wei, W X Zong, E H Cheng, T Lindsten, V Panoutsakopoulou, A J Ross, K A Roth, G R MacGregor, C B Thompson, and S J Korsmeyer. Proapoptotic bax and bak: a requisite gateway to mitochondrial dysfunction and death. *Science*, 292(5517):727–30, Apr 2001.

[582] Yongjie Wei, Sophie Pattingre, Sangita Sinha, Michael Bassik, and Beth Levine. Jnk1-mediated phosphorylation of bcl-2 regulates starvation-induced autophagy. *Mol Cell*, 30(6):678–88, Jun 2008.

[583] A Weiss and J Schlessinger. Switching signals on or off by receptor dimerization. *Cell*, 94(3):277–80, Aug 1998.

[584] A A Welihinda and R J Kaufman. The unfolded protein response pathway in saccharomyces cerevisiae. oligomerization and trans-phosphorylation of ire1p (ern1p) are required for kinase activation. *J Biol Chem*, 271:18181–7, Jul 1996.

[585] X-F Wen, G Yang, W Mao, A Thornton, J Liu, R C Bast, Jr, and X-F Le. Her2 signaling modulates the equilibrium between pro- and antiangiogenic factors via distinct pathways: implications for her2-targeted antibody therapy. *Oncogene*, 25(52):6986–96, Nov 2006.

[586] A J Whitmarsh and R J Davis. Transcription factor ap-1 regulation by mitogen-activated protein kinase signal transduction pathways. *J Mol Med*, 74(10):589–607, Oct 1996.

[587] Valérie Wilquet and Bart De Strooper. Amyloid-beta precursor protein processing in neurodegeneration. *Curr Opin Neurobiol*, 14(5):582–8, Oct 2004.

[588] Alejandro Wolf-Yadlin, Neil Kumar, Yi Zhang, Sampsa Hautaniemi, Muhammad Zaman, Hyung-Do Kim, Viara Grantcharova, Douglas A Lauffenburger, and Forest M White. Effects of her2 overexpression on cell signaling networks governing proliferation and migration. *Mol Syst Biol*, 2:54, 2006.

[589] Peter J Woolf, Wendy Prudhomme, Laurence Daheron, George Q Daley, and Douglas A Lauffenburger. Bayesian analysis of signaling networks governing embryonic stem cell fate decisions. *Bioinformatics*, 21(6):741–53, Mar 2005.

[590] Jun Wu, D Thomas Rutkowski, Meghan Dubois, Jayanth Swathirajan, Thomas Saunders, Junying Wang, Benbo Song, Grace D-Y Yau, and Randal J Kaufman. Atf6alpha optimizes long-term endoplasmic reticulum function to protect cells from chronic stress. *Dev Cell*, 13:351–64, Sep 2007.

[591] Xuesong Wu, Vivian C Lee, Eric Chevalier, and Sam T Hwang. Chemokine receptors as targets for cancer therapy. *Curr Pharm Des*, 15(7):742–57, 2009.

[592] Florian M Wurm. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol*, 22(11):1393–8, Nov 2004.

[593] Q Xu and J C Reed. Bax inhibitor-1, a mammalian apoptosis suppressor identified by functional screening in yeast. *Mol Cell*, 1(3):337–46, Feb 1998.

[594] Xiaojiang Xu, Lianshui Wang, and Dafu Ding. Learning module networks from genome-wide location and expression data. *FEBS Lett*, 578(3):297–304, Dec 2004.

[595] Takahiro Yamada, Hisamitsu Ishihara, Akira Tamura, Rui Takahashi, Suguru Yamaguchi, Daisuke Takei, Ai Tokita, Chihiro Satake, Fumi Tashiro, Hideki Katagiri, Hiroyuki Aburatani, Jun-ichi Miyazaki, and Yoshitomo Oka. Wfs1-deficiency increases endoplasmic reticulum stress, impairs cell cycle progression and triggers the apoptotic pathway specifically in pancreatic beta-cells. *Hum Mol Genet*, 15(10):1600–9, May 2006.

[596] K Yamamoto, H Ichijo, and S J Korsmeyer. Bcl-2 is phosphorylated and inactivated by an ask1/jun n-terminal protein kinase pathway normally activated at g(2)/m. *Mol Cell Biol*, 19:8469–78, Dec 1999.

[597] K. Yamamoto, H. Yoshida, K. Kokame, R.J. Kaufman, and K. Mori. Differential contributions of ATF6 and XBP1 to the activation of endoplasmic reticulum stress-responsive cis-acting elements ERSE, UPRE and ERSE-II. *Journal of biochemistry*, 136(3):343, 2004.

[598] Keisuke Yamamoto, Takashi Sato, Toshie Matsui, Masanori Sato, Tetsuya Okada, Hiderou Yoshida, Akihiro Harada, and Kazutoshi Mori. Tran-

scriptional induction of mammalian er quality control proteins is mediated by single or combined action of atf6alpha and xbp1. *Dev Cell*, 13:365–76, Sep 2007.

[599] Wei Yan, Christopher L Frank, Marcus J Korth, Bryce L Sopher, Isabel Novoa, David Ron, and Michael G Katze. Control of perk eif2alpha kinase activity by the endoplasmic reticulum stress-induced molecular chaperone p58ipk. *Proc Natl Acad Sci U S A*, 99:15920–5, Dec 2002.

[600] Jason H Yang and Jeffrey J Saucerman. Computational models reduce complexity and accelerate insight into cardiac signaling networks. *Circ Res*, 108(1):85–97, Jan 2011.

[601] K S Yao, S Xanthoudakis, T Curran, and P J O'Dwyer. Activation of ap-1 and of a nuclear redox factor, ref-1, in the response of ht29 colon cancer cells to hypoxia. *Mol Cell Biol*, 14(9):5997–6003, Sep 1994.

[602] Timothy A Yap, Michelle D Garrett, Mike I Walton, Florence Raynaud, Johann S de Bono, and Paul Workman. Targeting the PI3K-AKT-mTOR pathway: progress, pitfalls, and promises. *Curr Opin Pharmacol*, 8(4):393–412, Aug 2008.

[603] Y Yarden and M X Sliwkowski. Untangling the erbb signalling network. *Nat Rev Mol Cell Biol*, 2(2):127–37, Feb 2001.

[604] Terrence M Yau, Christopher Kim, Dawn Ng, Guangming Li, Yaoguang Zhang, Richard D Weisel, and Ren-Ke Li. Increasing transplanted cell survival with cell-based angiogenic gene therapy. *Ann Thorac Surg*, 80(5):1779–86, Nov 2005.

[605] J Ye, R B Rawson, R Komuro, X Chen, U P Davé, R Prywes, M S Brown, and J L Goldstein. Er stress induces cleavage of membrane-bound atf6 by the same proteases that process srebps. *Mol Cell*, 6:1355–64, Dec 2000.

[606] C H Yeang, S Ramaswamy, P Tamayo, S Mukherjee, R M Rifkin, M Angelo, M Reich, E Lander, J Mesirov, and T Golub. Molecular classification of multiple tumor types. *Bioinformatics*, 17 Suppl 1:S316–22, 2001.

[607] M K Stephen Yeung, Jesper Tegnér, and James J Collins. Reverse engineering gene networks using singular value decomposition and robust regression. *Proc Natl Acad Sci U S A*, 99(9):6163–8, Apr 2002.

[608] T Yoneda, K Imaizumi, K Oono, D Yui, F Gomi, T Katayama, and M Tohyama. Activation of caspase-12, an endoplastic reticulum (er) resident caspase, through tumor necrosis factor receptor-associated factor 2-dependent mechanism in response to the er stress. *J Biol Chem*, 276:13935–40, Apr 2001.

[609] Tomohiro Yorimitsu, Usha Nair, Zhifen Yang, and Daniel J Klionsky. Endoplasmic reticulum stress triggers autophagy. *J Biol Chem*, 281(40):30299–304, Oct 2006.

[610] H Yoshida, K Haze, H Yanagi, T Yura, and K Mori. Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose-regulated proteins. involvement of basic leucine zipper transcription factors. *J Biol Chem*, 273:33741–9, Dec 1998.

[611] H Yoshida, T Matsui, A Yamamoto, T Okada, and K Mori. Xbp1 mrna is induced by atf6 and spliced by ire1 in response to er stress to produce a highly active transcription factor. *Cell*, 107:881–91, Dec 2001.

[612] H Yoshida, T Okada, K Haze, H Yanagi, T Yura, M Negishi, and K Mori. Atf6 activated by proteolysis binds in the presence of nf-y (cbf) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Mol Cell Biol*, 20:6755–67, Sep 2000.

[613] Hiderou Yoshida. Er stress and diseases. *FEBS J*, 274:630–58, Feb 2007.

[614] Hiderou Yoshida, Masaya Oku, Mie Suzuki, and Kazutoshi Mori. pxbp1(u) encoded in xbp1 pre-mrna negatively regulates unfolded protein response activator pxbp1(s) in mammalian er stress response. *J Cell Biol*, 172:565–75, Feb 2006.

[615] Hiderou Yoshida, Aya Uemura, and Kazutoshi Mori. pxbp1(u), a negative regulator of the unfolded protein response activator pxbp1(s), targets atf6 but not atf4 in proteasome-mediated degradation. *Cell Struct Funct*, 34:1–10, 2009.

[616] D Yu and M C Hung. Overexpression of erbb2 in cancer and erbb2-targeting strategies. *Oncogene*, 19(53):6115–21, Dec 2000.

[617] Liang Yu, Lin Gao, and Peng Gang Sun. A hybrid clustering algorithm for identifying modules in protein-protein interaction networks. *Int J Data Min Bioinform*, 4(5):600–15, 2010.

[618] Wenli Yu, Louis-Bruno Ruest, and Kathy K H Svoboda. Regulation of epithelial-mesenchymal transition in palatal fusion. *Exp Biol Med (Maywood)*, 234(5):483–91, May 2009.

[619] T. L. Yuan and L. C. Cantley. PI3K pathway alterations in cancer: variations on a theme. *Oncogene*, 27(41):5497–5510, Sep 2008.

[620] Jiri Zavadil and Erwin P Böttinger. Tgf-beta and epithelial-to-mesenchymal transitions. *Oncogene*, 24(37):5764–74, Aug 2005.

[621] Q Zhan, K A Lord, I Alamo, Jr, M C Hollander, F Carrier, D Ron, K W Kohn, B Hoffman, D A Liebermann, and A J Fornace, Jr. The gadd and myd genes define a novel set of mammalian genes encoding acidic proteins that synergistically suppress cell growth. *Mol Cell Biol*, 14:2361–71, Apr 1994.

[622] C Zhang, J Kawauchi, M T Adachi, Y Hashimoto, S Oshiro, T Aso, and S Kitajima. Activation of jnk and transcriptional repressor atf3/lrf1 through the ire1/traf2 pathway is implicated in human vascular endothelial cell death by homocysteine. *Biochem Biophys Res Commun*, 289:718–24, Dec 2001.

[623] Haifeng Zhang, Rong Zhang, Yan Luo, Alessio D'Alessio, Jordan S Pober, and Wang Min. Aip1/dab2ip, a novel member of the ras-gap family, transduces traf2-induced ask1-jnk activation. *J Biol Chem*, 279:44955–65, Oct 2004.

[624] Hongtao Zhang, Mark Richter, and Mark I Greene. Therapeutic monoclonal antibodies for the erbb family of receptor tyrosine kinases. *Cancer Biol Ther*, 2(4 Suppl 1):S122–6, 2003.

[625] K Zhang and R J Kaufman. Protein folding in the endoplasmic reticulum and the unfolded protein response. *Handb Exp Pharmacol*, (172):69–91, 2006.

[626] K. Zhang and RJ Kaufman. Protein folding in the endoplasmic reticulum and the unfolded protein response. *Molecular chaperones in health and disease*, page 69, 2006.

[627] Kezhong Zhang and Randal J Kaufman. The unfolded protein response: a stress signaling pathway critical for health and disease. *Neurology*, 66(2 Suppl 1):S102–9, Jan 2006.

[628] Kezhong Zhang and Randal J Kaufman. From endoplasmic-reticulum stress to the inflammatory response. *Nature*, 454:455–62, Jul 2008.

[629] Kezhong Zhang, Xiaohua Shen, Jun Wu, Kenjiro Sakaki, Thomas Saunders, D Thomas Rutkowski, Sung Hoon Back, and Randal J Kaufman. Endoplasmic reticulum stress activates cleavage of crebh to induce a systemic inflammatory response. *Cell*, 124:587–99, Feb 2006.

[630] L Zhang, J Chen, and H Fu. Suppression of apoptosis signal-regulating kinase 1-induced cell death by 14-3-3 proteins. *Proc Natl Acad Sci U S A*, 96:8511 – 8515, 1999.

[631] Le Zhang, Chaitanya A Athale, and Thomas S Deisboeck. Development of a three-dimensional multiscale agent-based tumor model: simulating gene-protein interaction profiles, cell phenotypes and multicellular patterns in brain cancer. *J Theor Biol*, 244(1):96–107, Jan 2007.

[632] Wensheng Zhang, Andrea Edwards, Wei Fan, Dongxiao Zhu, and Kun Zhang. svdppcs: an effective singular value decomposition-based method for conserved and divergent co-expression gene module identification. *BMC Bioinformatics*, 11:338, 2010.

[633] Ying E Zhang. Non-smad pathways in tgf-beta signaling. *Cell Res*, 19(1):128–39, Jan 2009.

[634] Yanxiang Zhou, Juliane Liepe, Xia Sheng, Michael P H Stumpf, and Chris Barnes. Gpu accelerated biochemical network simulation. *Bioinformatics*, 27(6):874–6, Mar 2011.

[635] S. G. Zimmer, A. DeBenedetti, and J. R. Graff. Translational control of malignancy: the mRNA cap-binding protein, eIF-4E, as a central regulator of tumor formation, growth, invasion and metastasis. *Anticancer Res*, 20(3A):1343–1351, 2000.

[636] Roberto Zoncu, Alejo Efeyan, and David M Sabatini. mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol*, 12(1):21–35, Jan 2011.

[637] X Zou, T Tsutsui, D Ray, J F Blomquist, H Ichijo, D S Ucker, and H Kiyokawa. The cell cycle-regulatory cdc25a phosphatase inhibits apoptosis signal-regulating kinase 1. *Mol Cell Biol*, 21:4818–28, Jul 2001.